Methodology for Chemical Protein Synthesis

by

Lorraine Bland

A thesis submitted for the

degree of Doctor of Philosophy

University of Edinburgh

November 1999

This thesis is submitted in part fulfilment of the requirements for the degree of Doctor of Philosophy at the University of Edinburgh. Unless otherwise stated, the work is original and has not been previously submitted in whole or in part, for any degree at this, or any other university.

University of Edinburgh

November 1999

for mum, dad and Stephen

Acknowledgements

Firstly I would like to thank Professor R. Ramage, FRS for giving me the opportunity to study in Edinburgh and for his encouragement and advice throughout the course of the last three years.

Thanks also to the Edinburgh Centre for Protein Technology for financial support.

I would also like to thank Kevin Shaw and Brian Whigham for technical assistance with peptide synthesis, HPLC, amino acid analysis and MALDI ToF MS.

I am also extremely grateful to Dr Lesley Yellowlees for running the electrochemical analyses, Dr Simon Parsons for X-ray crystallographic work, Dr Andy Cronshaw (Wellmet Protein Characterisation Facility) for *N*-terminal sequencing and Mark Scott for assistance with electrospray mass spectrometry. Thanks also to the departmental technical staff for their efficient work.

Huge thanks go to all the members of lab 29, past and present, for their advice and friendship. A special mention has to go to Jennie Patterson, Craig Jamieson, Martin Andrews and Alastair Hay for helping me keep my sanity in the lab when things didn't go as planned. Also thanks to Stephen, Jennie and Martin for proof reading this thesis.

Finally I need to thank my mum, dad and Stephen for their endless love and support over the last three years, especially during the writing of this thesis, without which I would never have reached this stage.

Abstract

A number of groups have been investigated as potential protecting groups for the thiol functionality of cysteine residues during Solid Phase Peptide Synthesis. Of these the 4-picolyl system has been found to be successful. A short, efficient synthesis to the protected amino acid and mild cleavage conditions have been developed. This protected amino acid has been applied in the synthesis of a number of interesting cysteine-containing peptide fragments.

Deglycosylated human erythropoietin (dhEPO) has been synthesised by SPPS, employing the picolyl group for cysteine protection. Purification of this material has been achieved by gel filtration. The synthetic material has been characterised by amino acid analysis, SDS-PAGE, isoelectric focusing and *N*-terminal sequencing.

Attempts have been made to synthesise deglycosylated human erythropoietin *via* a fragment coupling technique. All fragments were successfully synthesised and purified and model ligation studies carried out.

Abbreviations

AAA	Amino Acid Analysis
Abs	absorbance
Acm	acetamidomethyl
AcOH	acetic acid
Allocam	Allyloxycarbonylaminomethyl
Boc	tertiary-butoxycarbonyl
br	broad
^{t}Bu	tertiary butyl
Bum	tertiary-butyloxymethyl
CHO	Chinese Hamster Ovary
COS-1	Green African Monkey Kidney Cells
d	doublet
Da	dalton
DBU	1,8-diazobicyclo[5.4.0]undec-7-ene
DCC	*N*, *N*-dicyclohexylcarbodiimide
DCM	dichloromethane
DIC	*N*, *N*-diisopropylcarbodiimide
DIEA	*N*, *N*-diisopropylethylamine
DMAP	4-dimethylaminopyridine
DMF	dimethylformamide
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
DTT	dithiothreitol
E. coli	*Escherichia coli*
EDT	ethanedithiol
EPO	erythropoietin
EPObp	erythropoietin binding protein
dhEPO	deglycosylated human erythropoietin
rhEPO	recombinant human erythropoietin

uhEPO	urinary human erythropoietin
EDTA	ethylenediaminetetraacetic acid
EDT	ethanedithiol
ES-MS	electrospray mass spectrometry
ET-1	endothelin-1
FAB	fast atom bombardment
Fmoc	9-fluorenylmethoxycarbonyl
FPLC	fast protein liquid chromatography
GRH	growth hormone
HF	hydrogen fluoride
HIV	human immunodeficiency virus
HOAt	1-hydroxy-7-azabenzotriazole
HOBt	1-hydroxybenzotriazole
HOCt	ethyl-1-hydroxy-1*H*-1,2,3-triazole-4-carboxylate
HPLC	high performance liquid chromatography
HRMS	high resolution mass spectrometry
Hz	hertz
IEF	isoelectric focusing
INF-γ	interferon gamma
IR	infrared
kDa	kilodalton
m	multiplet
MALDI	matrix assisted laser desorption ionisation
MHz	megahertz
MS	mass spectrometry
nm	nanometers
NMR	nuclear magnetic resonance
*i*NOC	isonicotinyloxycarbonyl
PAGE	polyacrylamide gel electrophoresis
PAM	hydroxymethylphenylacetamodimethyl
Pbf	2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl
PDI	Protein disulfide isomerase

Phacm	phenylacetamidomethyl
pI	isoelectric point
Pic	4-Picolyl
Pmc	2,2,5,7,8-pentamethylchroman-6-sulfonyl
p-TSA	para-toluene sulfonic acid
pTNM	4-methoxycarbonyl-4-nitro-[2,6-dioxaspiro-5:5-undecane]
ψ-Pro	Pseudo proline
q	quartet
R_t	retention time
s	singlet
SASRIN	super acid sensitive resin
SCD	catalytic domain of stromelysin
SDS	sodium dodecyl sulfonate
S^{t}Bu	tertiary butyl sulfenyl
SPOS	solid phase organic synthesis
SPPS	solid phase peptide synthesis
t	triplet
TAEC	transfer active ester condensation
Tbfmoc	tetrabenzo[*a,c,g,i*]fluorenyl-17-methoxycarbonyl
TBTU	2-(1H-benzotriazole-1-yl)-1,1,3,3-tetramethyl uroniumtetrafluoroborate
TFA	trifluoroacetic acid
TFMSA	trifluoromethanesulfonic acid
TIS	triisopropylsilane
Tlc	thin layer chromatography
TMS	tetramethylsilane
TNM	2,2-bishydroxymethyl-2-nitromethyloxycarbonyl
Tris	tris(hydroxymethyl)aminomethane
Trt	triphenylmethyl
UV	ultraviolet
Z	benzyloxycarbonyl

Contents

Chapter 1: Introduction to Solid Phase Peptide Synthesis

Chapter 2: Cysteine Protecting Groups

Appendix

CHAPTER 1

Introduction to Solid Phase Peptide Synthesis

1.1 Background

Peptides and proteins are primary metabolites found in all living cells. They are responsible for the maintenance and function of both the human and animal body and are the principal materials of skin, muscle, tendons, nerves, blood, enzymes, antibodies and many hormones. The word protein is derived from the Greek *proteios*, meaning first, indicating the importance with which these molecules were ranked even before their true purpose and function could be investigated.

The potential therapeutic use of these natural compounds is huge and already a number of peptides are manufactured on a kilogram scale for clinical use. For example Zoladex,[1] which is manufactured by Zeneca for the treatment of breast and prostate cancer, and oxytocin, which is often administered during labour. Moreover proteins produced by recombinant DNA technology have also been licensed for medical purposes. Examples here include human erythropoietin,[2] used for the treatment of anaemia associated with kidney disease, and interferon-β,[3] used in the treatment of multiple sclerosis. Isolation of peptides and proteins from natural sources is possible although yields are often low. Analysis of this material can provide the primary sequence of amino acids in the protein backbone. With this information there are two routes available to increase the quantity of material available. Using recombinant technology the gene, coding for the protein, can be cloned and used to express the protein in mammalian cells. Alternatively, a purely synthetic route can be adopted to form the linear sequence of amino acids.

A great deal of research has been invested into obtaining proteins *via* recombinant techniques, however, there are drawbacks to this method. Expression levels of the protein can be low and, perhaps more seriously, point mutations in the amino acid sequence can occur resulting in expression of the wrong material. This problem is

eliminated in chemical protein synthesis since the chemist has complete control over the sequence being synthesised. Single point mutations can have disastrous effects. For example, when a glutamate residue in haemoglobin is substituted for valine sticky patches develop on the protein surface, causing red blood cells to adopt the 'sickle' shape characteristic of the disease sickle cell anemia (**Figure 1.1**).[4] As a result transportation of red blood cells around the body is hindered leading to a potentially fatal condition.

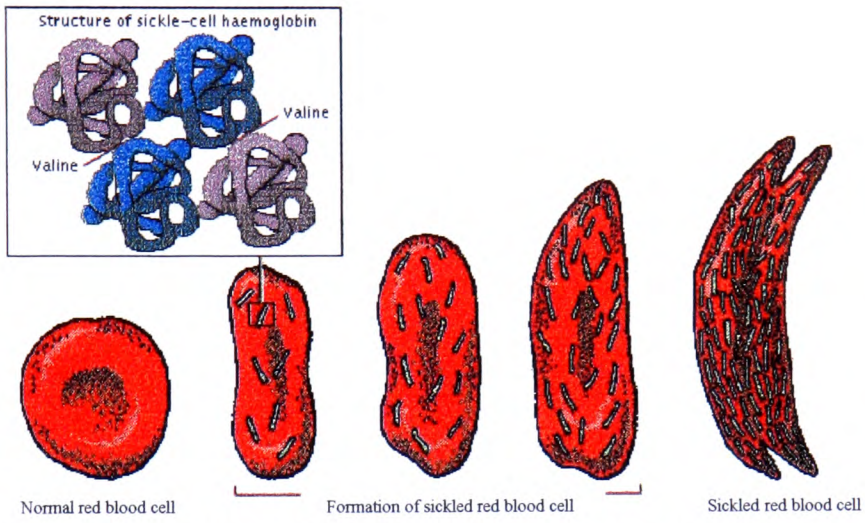


Figure 1.1 Mutation Leading to Sickle Cell Anemia

1.2 Peptide Synthesis

The chemical synthesis of peptides has been a challenge addressed for many years. The laborious task of solution phase peptide synthesis can be highlighted with the synthesis of oxytocin,[5,6] for which Du Vigneaud received the Nobel Prize in 1955. Although the synthesis of such a small peptide may appear trivial, the unambiguous formation of a dipeptide from two different amino acids requires careful consideration of the likely impurities. There are four possible products from this reaction, as outlined in **Figure 1.2**. Also, there is the risk of unwanted side reactions occurring at the side chain functional groups of the amino acids.

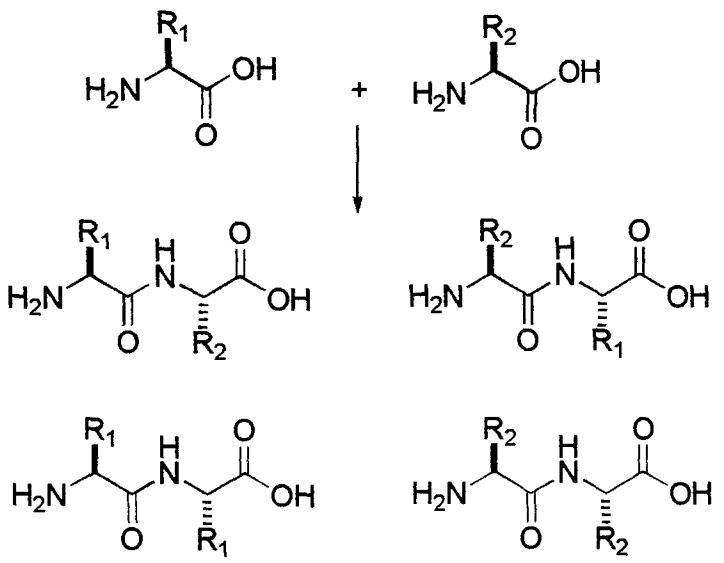


Figure 1.2 Products from Uncontrolled Dipeptide Synthesis

By blocking the carboxy group of one amino acid and the amino group of the second, activation of the free carboxy group enables directed formation of the peptide bond. Protection of any side chain functional groups prevents side reactions occurring. It is this protecting group strategy which forms the basis of contemporary peptide synthesis.

1.3 Solid Phase Peptide Synthesis

Merrifield[7] revolutionised peptide synthesis with the introduction of solid phase peptide synthesis (SPPS), a fact recognised with the award of the Nobel Prize for chemistry in 1984.[8] The basic concept of SPPS is outlined in **Figure 1.3**. It involves anchoring the first N^{α} protected amino acid to an insoluble solid support *via* a covalent bond. Cleavage of the N^{α} protecting group and subsequent coupling of the next amino acid can be repeated in a stepwise manner, generating the desired peptide sequence. Excess reagents are used to drive the reactions to completion and then simply washed away from the solid support. On completion of the synthesis the peptide is simply cleaved from the solid support and purified. Thus the rigorous purification of each intermediate peptide required *via* conventional solution phase methods is avoided. From the outset it was noted that the method should be adaptable to automation.

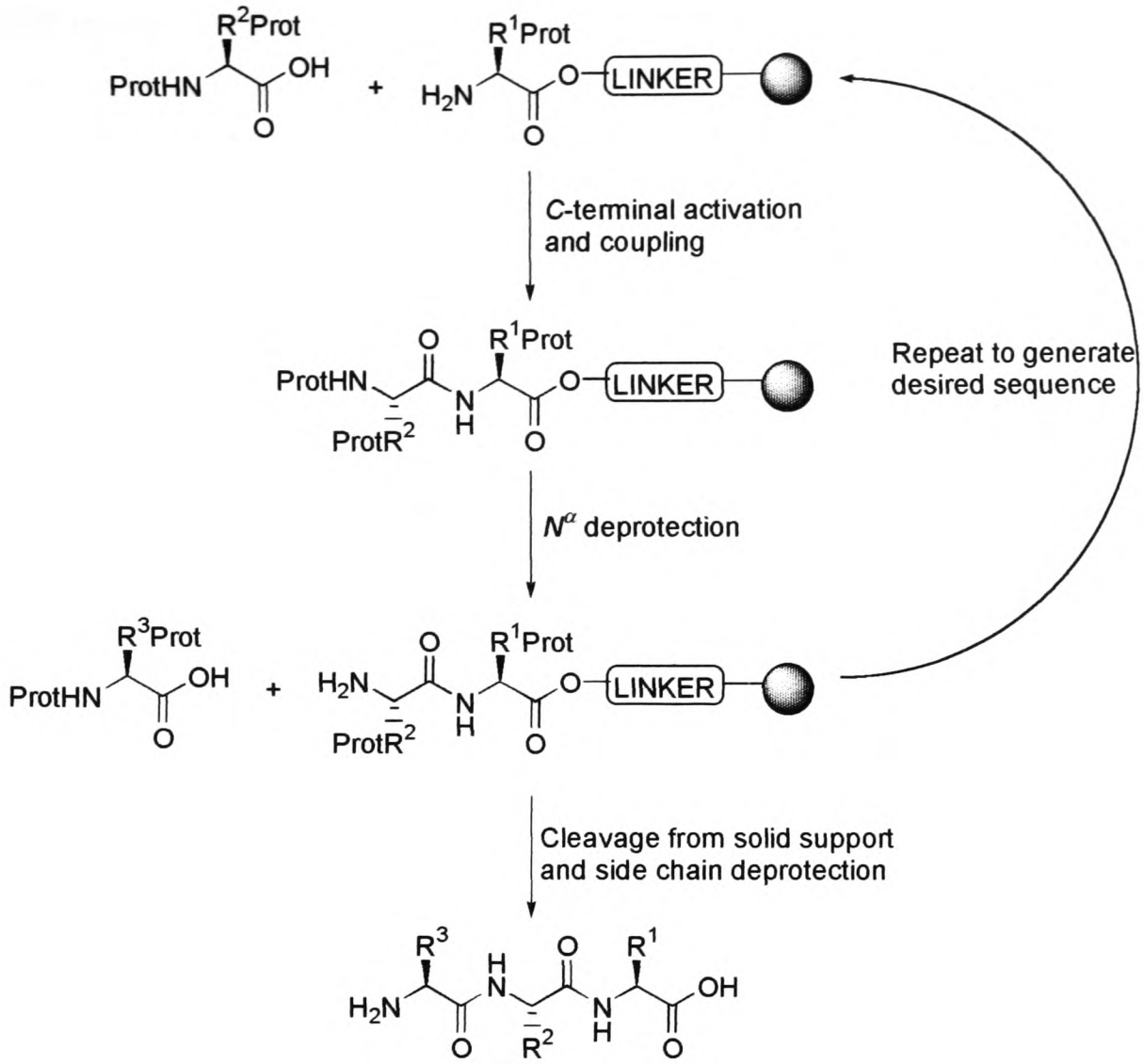


Figure 1.3 General Outline of SPPS

1.4 The Solid Support

The novel concept introduced by Merrifield was that of the solid support. If an insoluble polymer could be developed to swell in organic solvents a solvated gel would be produced. Reagents could then penetrate within the matrix of this gel enabling reactions to take place. By using excess reagents, the reactions could be driven to completion. Obviously the resin would need to possess a sufficiently rigid structure to enable the excess reagents to be removed by a simple filtration step. Functionalisation of the resin would facilitate covalent attachment of the first residue. With the increasing length of the syntheses, the stability of the resin was also crucial, considering the number of chemical steps to which it would be exposed.

Merrifield investigated a number of polymeric supports and found the best to be a copolymer of styrene crosslinked with 1% divinylbenzene (**Figure 1.4**).[7] Swelling of the resin was observed in DCM and DMF.

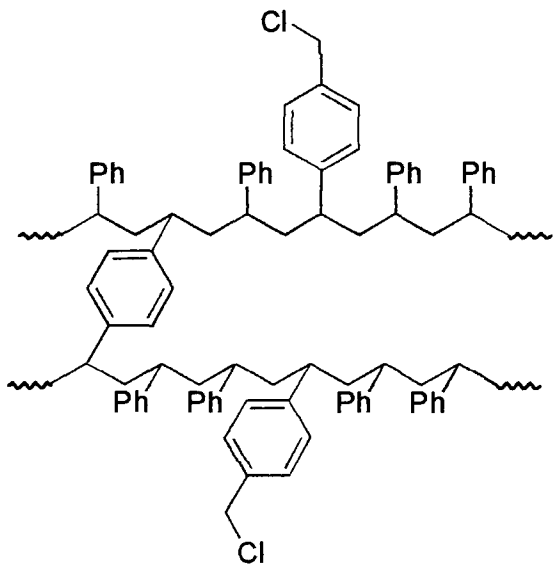


Figure 1.4 Polystyrene Crosslinked with 1% Divinyl Benzene

Although this support is still widely used in SPPS a number of others have been proposed. The most notable of these are the polyacrylamide resins developed by Sheppard and co-workers (**Figure 1.5**).[9,10,11] These are much less hydrophobic than the polystyrene based resins and resemble the polarity of the growing polypeptide chain. As a result, both the resin and the peptide chain are solvated. This appears to help in the synthesis of peptides which tend to aggregate on the polystyrene supports. These resins are compatible with both Fmoc and Boc methodology, which will be discussed in **Section 1.6**.

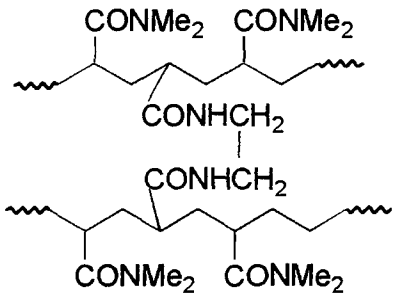


Figure 1.5 Polyacrylamide Resin

Tentagel®[12] and PEGA[13] resins have been developed for use in SPPS. PEGA swells extensively in a range of organic solvents and is freely permeable to macromolecules of up to 35 kDa, making it ideal for carrying out on-resin enzyme assays. Tentagel®

has excellent swelling properties in a number of solvents, including water and methanol. Both resins are able to maintain good flow rates in columns, making them particularly useful for continuous flow SPPS.

1.5 Linkers

In order to carry out any chemistry on the solid support, the resin must be functionalised to enable covalent attachment of the first residue. Modification of the *C*-terminal functionality and the cleavage conditions for the peptide–resin bond can be controlled in this way. Merrifield originally chloromethylated the polystyrene resin, enabling attachment of the first amino acid as the benzyl ester.[7] The harsh acidic conditions required for removal of the N^{α} protecting group, benzyloxycarbonyl (Z), did, however, cause slight loss of the peptide from the resin through attack of the benzyl ester linkage, **Figure 1.6**. Obviously this problem became more serious with the increasing length of syntheses. Nitration or bromination of the resin helped minimise this problem.

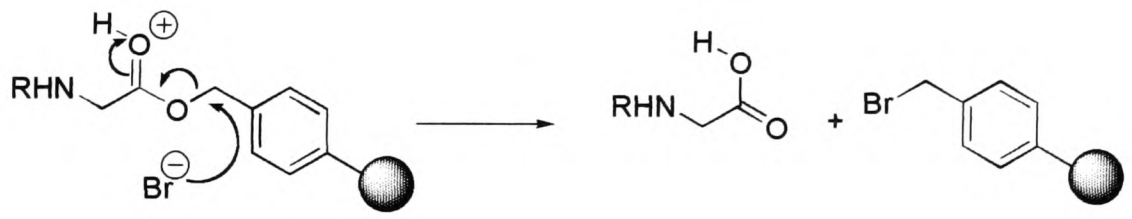


Figure 1.6 Cleavage of the Benzyl Ester Linkage

The introduction of the hydroxymethylphenylacetamidomethyl (PAM) linker increased the stability of the anchoring bond. Cleavage is accomplished using hydrogen fluoride (HF).[14]

The introduction of the base labile 9-fluorenylmethoxycarbonyl (Fmoc) group for N^{α} protection, led to the development of linkers which could be cleaved under much milder acidic conditions. The alkoxybenzylalcohol (Wang) linker is readily cleaved by trifluoroacetic acid (TFA), eliminating the need for HF or

trifluoromethanesulfonic acid (TFMSA), and has found wide spread use in Fmoc SPPS.

Advances in peptide synthesis, combined with the recent explosion of solid phase organic synthesis (SPOS), have lead to the introduction of a variety of new linkers, which give greater control over the chemistry being performed. The most useful of these, which are encountered in SPPS, are outlined in **Table 1.1**.

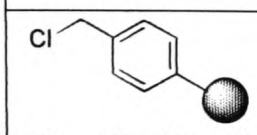
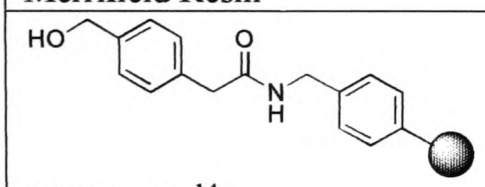
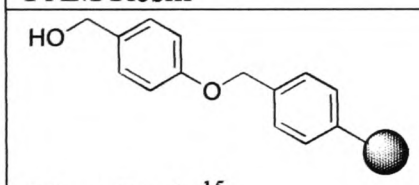
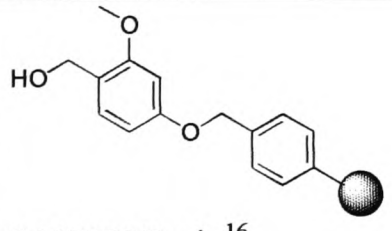
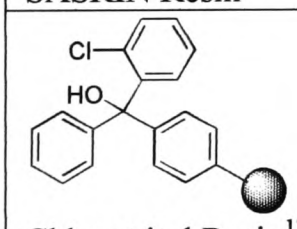
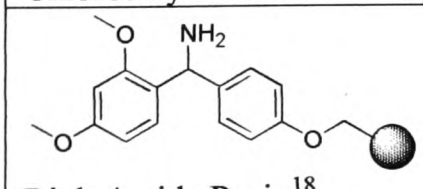
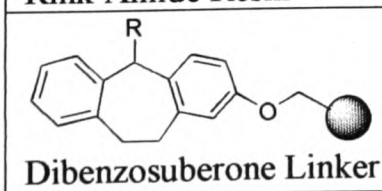
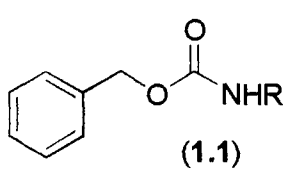
Linker	**Cleavage Conditions of Corresponding Ester**
 Merrifield Resin[7]	Original chloromethylated polystyrene resin used by Merrifield. Cleavage with HBr in glacial acetic acid.
 PAM Resin[14]	Cleavage with HF or TFMSA. Resin of choice for Boc SPPS.
 Wang Resin[15]	Cleavage with 95 % TFA / water. Resin of choice for Fmoc SPPS.
 SASRIN Resin[16]	Variant of the Wang linker. Cleavage occurs under very mild conditions – typically 0.5 – 1 % TFA in DCM.
 Chlorotrityl Resin[17]	Very acid labile – can cleave with 10% Acetic acid / water. Useful for the preparation of side chain protected peptide fragments.
 Rink Amide Resin[18]	Cleavage with dilute TFA generates the peptide amide.
 Dibenzosuberone Linker[19]	Cleavage with aqueous TFA. By varying the nature of the R group can generate a variety of peptide *C*-terminal derivatives.

Table 1.1 Common Linkers Encountered in SPPS

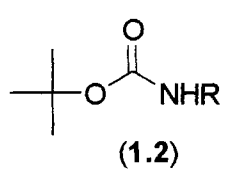
1.6 Protecting Group Strategies

1.6.1 Boc Strategy

Initially Merrifield employed the *Z* group (**1.1**) for N^{α} protection, the problems associated with which have been detailed earlier (**Section 1.5**). In a subsequent paper, detailing the synthesis of bradykinin,[20] Merrifield employed the more acid labile *t*-butyloxycarbonyl (Boc)[21,22] group (**1.2**) for N^{α} protection. Quantitative cleavage could be accomplished using milder acids such as TFA, enabling protection of the side chain functional groups with the Z group and as benzyl ethers and esters. Side chain deprotection and cleavage of the benzyl ester linker could be achieved at the end of the synthesis using HF or TFMSA. This work still forms the basis for Boc SPPS as carried out today, although there are concerns that the repetitive TFA treatment, required to cleave the Boc group, and the final cleavage with HF, may be destructive to some peptide sequences.[11]



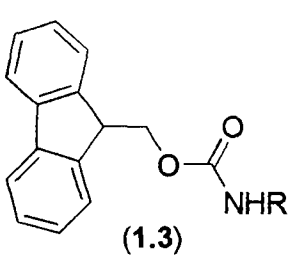
(1.1)



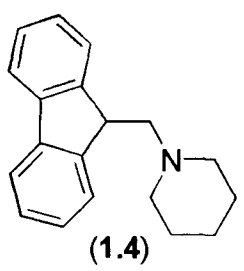
(1.2)

1.6.2 Fmoc Strategy

The truly orthogonal approach to SPPS was realised with the development of the base labile 9-fluorenylmethoxycarbonyl (Fmoc) group (**1.3**) by Carpino and Han.[23,24] This could be used for N^{α} protection in conjunction with ^{t}Bu derived acid labile side chain protecting groups, thus eliminating the use of HF or TFMSA. Cleavage of the Fmoc group occurs swiftly with secondary amines such as piperidine. Cleavage of the side chain protecting groups and peptide resin linkage can then be achieved using TFA. Another advantage encountered with this methodology is that the fulvene - piperidine adduct (**1.4**) formed during cleavage has a UV-maximum at 302nm.



(1.3)



(1.4)

This provides a useful means of monitoring the coupling efficiency of each step of the synthesis.[25] As a result, problematic areas in the synthesis can be identified and steps taken to rectify these in any subsequent syntheses.

1.6.3 Side Chain Protecting Groups

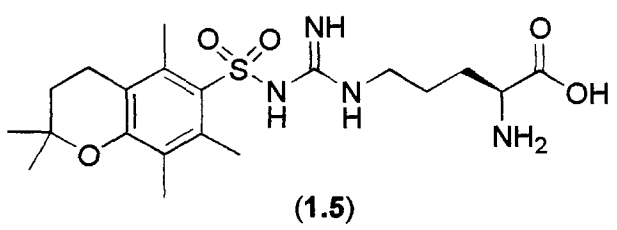
Amino acids contain a variety of side chain functional groups. These require protection during SPPS to prevent unwanted side reactions. Ideally these protecting groups should be stable to all the conditions employed during SPPS yet be readily cleaved at the end of the synthesis. As mentioned above, when the Boc group is used for N^{α} protection, the side chain functional groups can be protected with the Z group, benzyl ethers and esters. This approach relies on the greater acid lability of Boc over these groups.

With the introduction of the Fmoc group for N^{α} protection, a new approach to side chain protection was developed. The base lability of the Fmoc group meant that side chain functionality could now be protected with groups which could be cleaved under much milder acidic conditions. The most commonly used groups are triphenylmethyl (Trt), *t*-butyl ethers and *t*-butyl esters as outlined in **Table 1.2**.

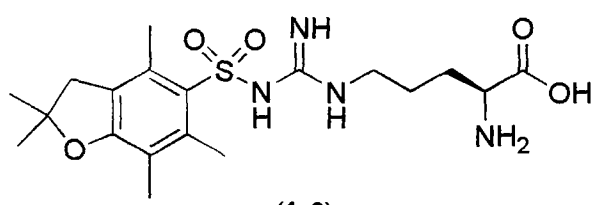
Amino Acid	Side Chain Protecting Group
Arg	Pmc / Pbf
Asn / Gln	Trt
Asp / Glu	^{t}Bu esters
Cys	Trt / Acm / ^{t}Bu / S^{t}Bu / Phacm
His	π-Bum / τ-Trt
Lys	Boc
Ser / Thr / Tyr	^{t}Bu ethers
Trp	Boc

Table 1.2 Side Chain Protecting Groups Employed in Fmoc SPPS

Protection of cysteine, arginine, tryptophan and histidine requires more complicated methodology. Due to the extremely nucleophilic nature of the guanidino function of arginine the Boc group is not sufficient to prevent side reactions at these sites. This problem has been addressed with the development of the 2,2,5,7,8-pentamethylchroman-6-sulfonyl (Pmc) group (**1.5**) which can be cleaved with 95% TFA in 1 hour.[26] It has been noted, however, that in peptides containing multiple arginine residues, the cleavage time must be increased to up to 5 hours to ensure complete removal. Slight modification of this group resulted in the development of the 2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl (Pbf) group (**1.6**) which is found to cleave slightly faster in 95% TFA.[27]

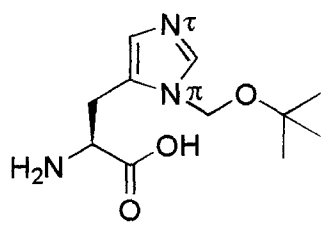


(1.5)

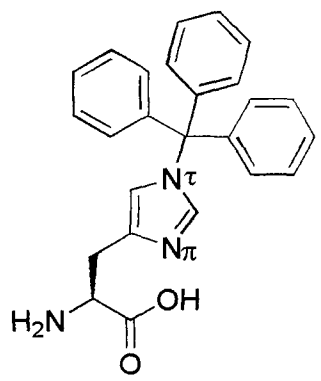


(1.6)

Histidine is prone to undergo racemisation due to the basicity of the π-nitrogen in the imidazole ring, which may remove the α-proton producing a planar intermediate. Racemisation can be completely eliminated by blocking the π-nitrogen with the *t*-butyloxymethyl (Bum) group (**1.7**).[28] Unfortunately, the multi-step synthesis required for production of this group, renders it a financially unviable option. Blocking the τ-nitrogen with the trityl group (**1.8**) reduces the bacisity of the π-nitrogen, thus suppressing racemisation.



(1.7)



(1.8)

The indole ring of tryptophan does not interfere with peptide bond formation, but presents a problem at the cleavage step since it captures electrophiles with great ease. As a result, scavengers are essential to sequester the carbocations formed from

cleavage of protecting groups. This problem can be partially addressed with the use of competitive scavengers during the cleavage step and also by protection of the nitrogen of the indole ring with the Boc group (**1.9**).[29] Cleavage with TFA generates an *N*-carboxy indole (**1.10**), which is subsequently cleaved under the aqueous conditions employed during work up of the peptide (**Figure 1.7**). This approach is particularly effective when used in conjunction with thiol and silane containing scavengers.

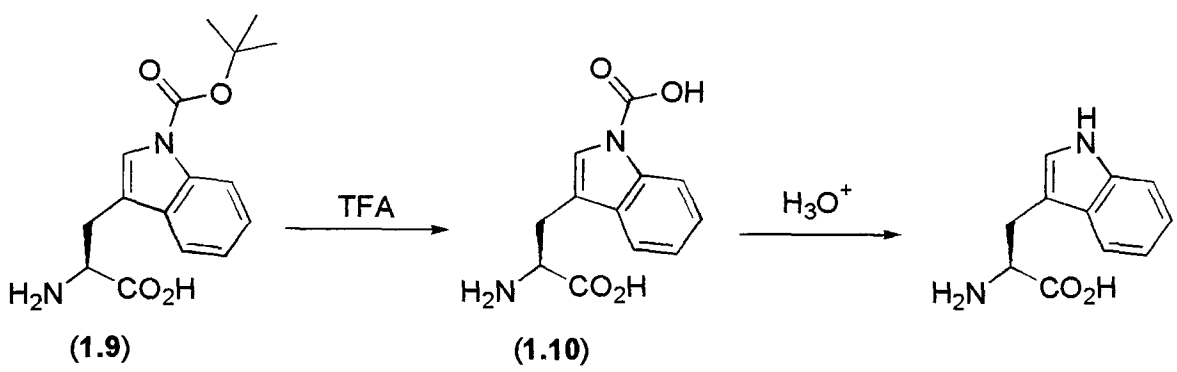


Figure 1.7 Protection of Tryptophan with the Boc Group

Cysteine containing peptides often require complicated methodology. This area will be reviewed in **Chapter 2**.

1.7 Peptide Bond Formation

The principal reaction in peptide synthesis is the formation of the amide bond between two amino acids. It is essential that this reaction be fast, high yielding and free of racemisation. This reaction proceeds most efficiently when the carbonyl group is activated, producing a more reactive intermediate. Care must be taken to avoid strongly electron withdrawing activating groups since these can induce racemisation. Racemisation is base induced occurring *via* an oxazolone intermediate (**Figure 1.8**). Urethane derived protecting groups such as Boc and Fmoc help suppress this unwanted reaction,[30] although it can still proceed under strongly basic conditions.[31]

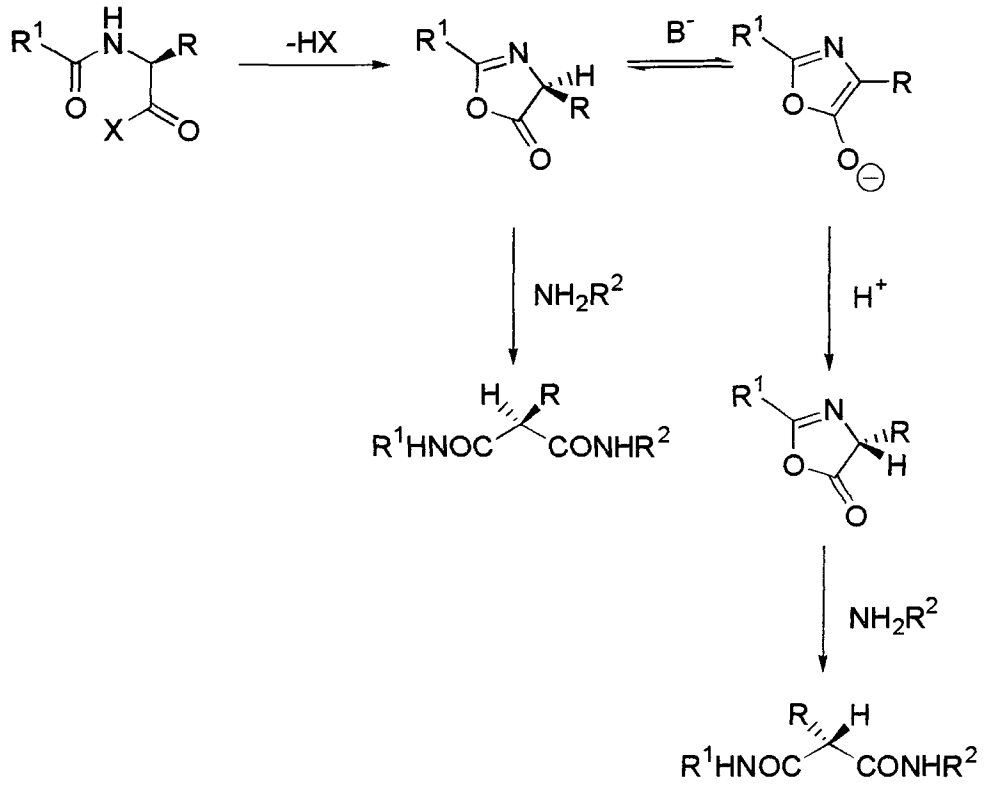


Figure 1.8 Racemisation via the oxazolone intermediate

1.7.1 Carbodiimides

Dicyclohexylcarbodiimide (DCC) (**1.11**) was introduced as an activating agent for the formation of amide bonds in 1955.[32] Merrifield demonstrated the synthetic utility of DCC as an activating agent in SPPS in his original paper.[7] Diisopropylcarbodiimide (DIC) (**1.12**)[33] has now become the reagent of choice in SPPS. This is simply due to the enhanced solubility of the *N,N'*-diisopropylurea by-product in the solvents used in SPPS, enabling easier separation from the solid support.

Carbodiimides react with the free carboxyl group of an amino acid forming an *O*-acyl urea (**1.13**). This can then go on to react in one of three ways (**Figure 1.9**);

1. React directly with the free amino function to generate the peptide bond (**1.14**).
2. Form a symmetrical anhydride (**1.15**) from reaction with a second equivalent of the amino acid.
3. React with another oxygen nucleophile to form a more stable intermediate such as an active ester (**1.16**).

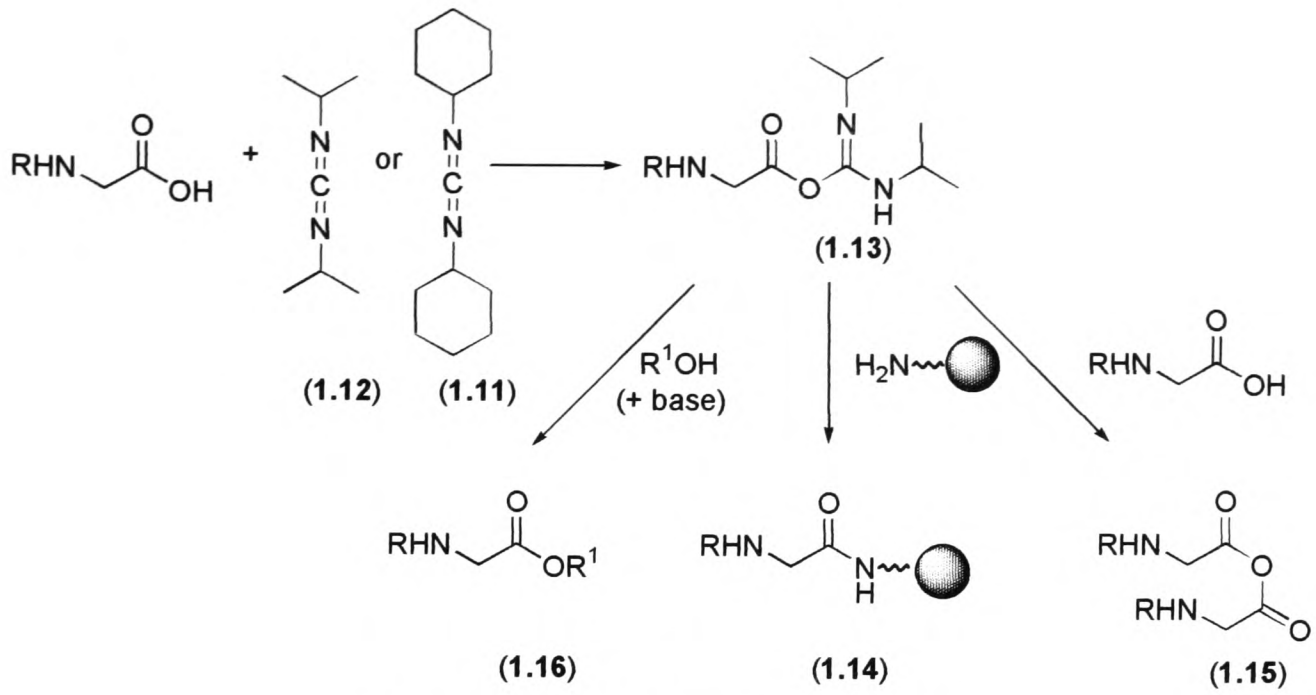


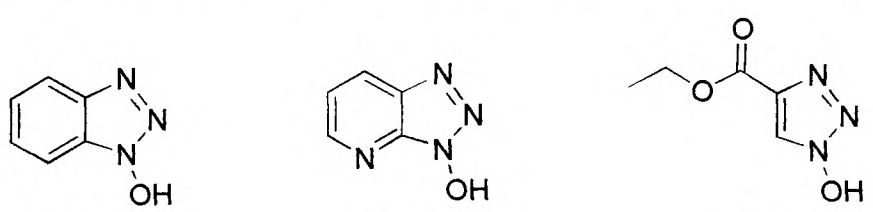
Figure 1.9 Activation of Amino Acids

1.7.2 Symmetrical Anhydrides

As mentioned above, symmetrical anhydrides are formed when the *O*-acyl urea reacts with a second equivalent of the amino acid. These reagents have found great utility in peptide synthesis, due to their ease of preparation and excellent acylating properties. There is, however, a drawback since two equivalents of the amino acid are required. Half of the amino acid is then essentially wasted and so this becomes a very expensive method of peptide bond formation.

1.7.3 Active Esters

The potential of active esters was first noted in 1964 when *p*-nitrophenol esters were coupled under modified SPPS conditions.[34] 1-Hydroxybenzotriazole (HOBt) (**1.17**) was introduced as a co-reagent with carbodiimide to prevent racemisation and *N*-acyl urea formation.[35] The success achieved by this method resulted in the preparation of analogues of HOBt in an attempt to improve on coupling efficiency. Both 1-hydroxy-7-azabenzotriazole (HOAt) (**1.18**)[36] and ethyl-1-hydroxy-1H-1,2,3-triazole-4-carboxylate (HOCt) (**1.19**)[37] have been found to be excellent coupling reagents.



HOCt has been applied successfully in the stepwise synthesis of the large proteins EPO[38], IFN-γ[39,40] and MMP-3[41], which were unobtainable using HOBt mediated coupling. More favourably, the unwanted side reaction, causing irreversible nitrile formation of the side chain of asparagine and glutamine residues (**Figure 1.10**),[42] often prevalent with carbodiimides and symmetrical anhydrides is not observed with active esters.

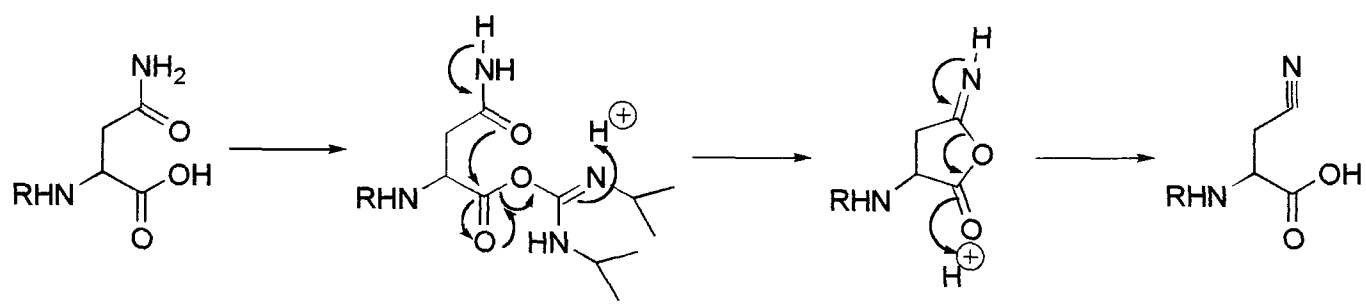


Figure 1.10 Dehydration of the Asparagine Side-chain

Extensive racemisation studies have been carried on HOCt mediated couplings. Of the twenty naturally occurring amino acids, only histidine was shown to racemise, a problem encountered with a number of activating agents. This could be reduced to a negligible level with the introduction of a three fold molar excess of HOCt.[43]

1.8 Advances in Peptide Synthesis

Changes in methodology are required in order to access large proteins of greater than 200 residues. Unless each step in SPPS is 100% efficient, there is a limit to which the methodology can stretch. By breaking a protein into smaller peptide fragments (typically 50 amino acids), it can be envisioned that these could then be readily synthesised and purified. Ligation of these smaller fragments could then generate the desired sequence.

It is nearly 100 years since the azide method for peptide bond formation was introduced by Curtius.[44] This methodology was adapted by Hirschmann to synthesise ribonuclease A, a 104 residue protein, *via* fragment coupling.[45]

Recently Kent has introduced a new method called native chemical ligation.[46] The initial work involved the reaction of a *C*-terminal α-thiocarboxylate on one fragment with a *N*-terminal α-bromoacyl moiety of a second fragment. The resultant ligated material having a thioester bond at the ligation site. The mild acidic conditions required to carry out the reactions render all side chain functional groups on the protein backbone non-nucleophilic, thus side chain protection is not required. The methodology has since been extended to generate a natural amide bond at the ligation site.[47] Here a *C*-terminal α-thioester is reacted with an *N*-terminal cysteine residue, generating the thioester intermediate (**1.20**), which then undergoes spontaneous rearrangement to give the native amide bond and the free thiol side chain of the cysteine residue (**Figure 1.11**). Further modification of this work means that it is no longer limited to *N*-terminal cysteine residues.[48]

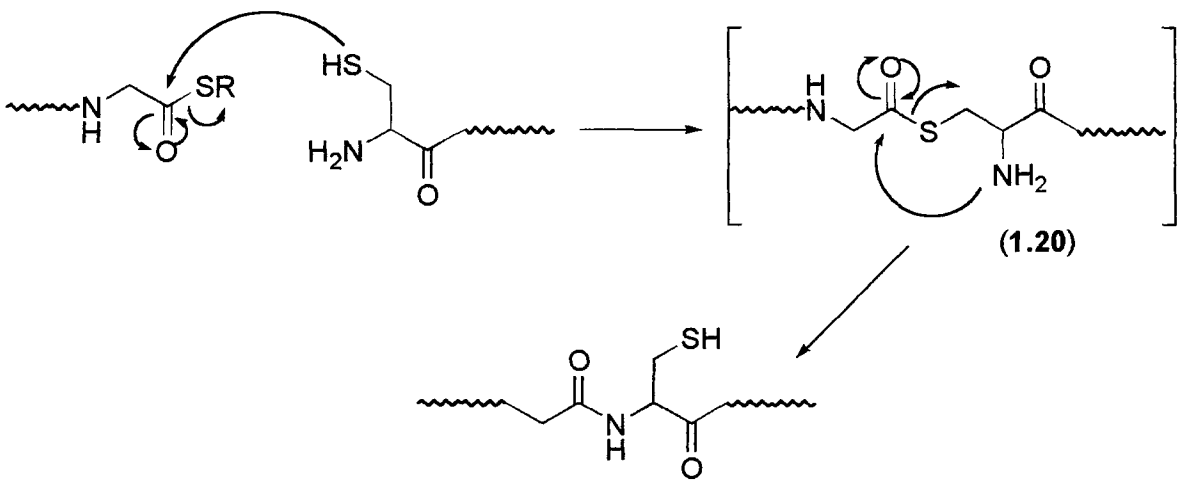


Figure 1.11 General Outline of Native Chemical Ligation

1. A. S. Dutta, US Patent 4,100,274

2. F. K. Lin, US Patent 4,703,008

3. G. L. Hall, A. Compston and N. J. Scolding, *Trends Neurosci.*, 1997, **20**, 63.

4. "Sickle-Cell Anaemia," *Microsoft® Encarta® 98 Encyclopedia.* © 1993-1997 Microsoft Corporation. All rights reserved.

5. V. Du Vigneaud, C. Ressler, J. M. Swan, C. W. Robertsl, P. G. Katsoyannis and S. Gordon, *J. Am. Chem. Soc.*, 1953, **75**, 4879.

6. V. Du Vigneaud, C. Ressler, J. M. Swan, C. W. Roberts and P. G. Katsoyannis, *J. Am. Chem. Soc.*, 1954, **76**, 3115.

7. R. B. Merrifield, *J. Am. Chem. Soc.*, 1963, **85**, 2149.

8. R. B. Merrifield, *Science*, 1986, **232**, 341.

9. E. Atherton, D. L. J. Clive and R. C. Sheppard, *J. Am. Chem. Soc.*, 1975, **97**, 6585.

10. R. Arshady, E. Atherton, D. L.J. Clive & R. C. Sheppard, *J. Chem. Soc., Perkin Trans. I.*, 1981, 529.

11. E. Atherton, C. J. Logan and R. C. Sheppard, *J. Chem. Soc., Perkin Trans. I*, 1981, 538.

12. W. Rapp, L. Zhang & E. Bayer, in *"Solid Phase Syntheis" (R. Epton, Ed.) SPPC, Birmingham, UK*, 1990, 205.

13. M. Medal, *Tetrahedron Lett.*, 1992, **33**, 3077.

14. A. R. Mitchell, S. B. H. Kent, M. Engelhard and R. B. Merrifield, *J. Org. Chem.*, 1978, **43**, 2845.

15. S. S. Wang, *J. Am. Chem. Soc.*, 1973, **95**, 1327.

16. M. Mergler, R. Tanner, J. Gosteli and P. Grogg, *Tetrahedron Lett.*, 1988, **29**, 4005.

17. K. Barlos, D. Gatos, J. Kallitsis, G. Papaphotiu, P. Sotiriu, Y. Wenging and W. Schäfer, *Tetrahedron Lett.*, 1989, **30**, 3943.

18. H. Rink, *Tetrahedron Lett.*, 1987, **28**, 3787.

19. R. Ramage, S. L. Irving and C. M^{c}Innes, *Tetrahedron Lett.*, 1993, **34**, 6599.

20. R. B. Merrifield, *Biochemistry*, 1964, **3**, 1385.

21. L. A. Carpino, *J. Am. Chem. Soc.*, 1957, **79**, 98.

22. L. A. Carpino, *J. Am. Chem. Soc.*, 1957, **79**, 4427.

23. L. A. Carpino and G. Y. Han, *J. Am. Chem. Soc.*, 1970, **92**, 5748.

24. L. A. Carpino and G. Y. Han, *J. Org. Chem.*, 1972, **37**, 3404.

25. K. M. Otteson, R. L. Nobel, P. D. Hoeprich Jr., K. T. Shaw and R. Ramage, *Applied Biosystems Res. News*, June 1993.

26. R. Ramage and J. Green, *Tetrahedron Lett.*, 1987, **28**, 2287.

27. L. A. Carpino, H. Shroff, S. A. Triolo, E. M. E. Manisour, H. Wenschuh and F. Albericio, *Tetrahedron Lett.*, 1993, **34**, 7829.

28. R. Colombo, F. Colombo and J. H. Jones, *J. Chem. Soc. Chem. Commun.*, 1984, 292.

29. P. White in *'Peptides, Chemistry and Biology, Proc., 12th US Peptide Symposium'*, J. A. Smith and J. E. River (Eds), ESCOM, Leiden, 1992, p 537.

30. D. S. Kemp, in *'The Peptides'*, Volume I, 1979, Academic Press Inc., pp 315 – 383.

31. R. Colombo, E. Atherton, R. C. Sheppard and V. Woollery, *Int. J. Peptide Protein Res.*, 1983, **21**, 118.

32. J. C. Sheehan and G. P. Hess, *J. Am. Chem. Soc.*, 1955, **77**, 1067.

33. D. Sarantakis, J. Teichman, E. L. Lien and R. L. Fenichel, *Biochem. Biophys. Res. Commun.*, 1976, **73**, 336.

34. M. Bodansky and J. T. Sheehan, *J. Chem. Ind. (London)*, 1964, **32**, 1432.

35. W. König and R. Geiger, *Chem. Ber.*, 1970, **103**, 788.

36. L. A. Carpino, *J. Am. Chem. Soc.*, 1993, **115**, 4397.

37. L. Jiang, A. Davison, G. Tennant and R. Ramage, *Tetrahedron*, 1998, **54**, 14233.

38. N. Robertson and R. Ramage, *J. Chem Soc., Perkin Trans. I*, 1999, 1015.

39. L. C. Draffan, *Ph. D. Thesis*, University of Edinburgh, 1996.

40. J. A. Patterson, *Ph. D. Thesis*, University of Edinburgh, 1999.

41. G. H. Morton, *Ph. D. Thesis*, University of Edinburgh, 1997.

42. J Blake and C. H. Li, *Int. J. Pept. Prot. Res.*, 1975, 7, 495.

43. N. Robertson, L. Jiang and R. Ramage, *Tetrahedron*, 1999, **55**, 2713.

44. T. Curtius, *Ber. Dtsch. Chem. Ger.*, 1902, **35**, 3226.

45. R. Hirschmann, R. F. Nutt, D. F. Veber, R. A. Vitali, S. L. Varga, T. A. Jacob, F. W. Holly and R. G. Denkewalter, *J. Am. Chem. Soc.*, 1969, **91**, 507.

46. M. Schnöler and S. B. H. Kent, *Science*, 1992, **256**, 221.

47. P. E. Dawson, T. W. Muir, I. Clark-Lewis and S. B. H. Kent, *Science*, 1994, **266**, 776.

48. L. E. Canne, S. J. Bark and S. B. H. Kent, *J. Am. Chem. Soc.*, 1996, **118**, 5891.

CHAPTER 2
Cysteine Protecting Groups

2.1 Background

Once the linear sequence of amino acids has been formed, a peptide or protein must fold into its biologically active form before it can produce any effect *in vivo* or *in vitro*. Secondary structure forms with the development of α-helices, β-sheets and β-turns, through hydrogen bonding and ionic and hydrophobic interactions between the side chains of the amino acids. However, it is the development of tertiary structure which defines the total three-dimensional conformation of the polypeptide chain. The thiol side chains of cysteine residues play a key role in the development of this tertiary structure, since the oxidative linking of two cysteine residues generates a disulfide.

As mentioned previously, it is necessary to temporarily mask the side chain functionality of amino acids during SPPS, to prevent side reactions occurring at these sites. Cysteine is no exception to this rule, the reactivity of the thiol side chain necessitating its protection. Indeed, it can be argued that purification of chemically synthesised peptides containing cysteine can be simplified if the thiol remains protected during this process.

2.2 Orthogonal Protecting Groups and Disulfide Formation

In proteins and peptides containing two cysteine residues, disulfide bond formation can be carried out quite simply by *S*-deprotection followed by oxidation. Peptides or proteins containing two or more disulfide bridges require more complicated methodology since the *S*-*S* bond formation must be directed to produce the correctly folded material. Essentially there are three approaches to this problem;

1. Subject the fully deprotected material to folding conditions and identify the correct material *via* comparison with a standard or by structural analysis.
2. Coupling of fragments with pre-formed disulfide bridges. This method is highly sequence dependant.
3. Directed disulfide formation through the use of orthogonal cysteine protecting groups.

Of these methods the use of orthogonal protecting groups is possibly the most effective. For this reason a number of protecting groups have been proposed for cysteine residues in SPPS. These will be reviewed in **Section 2.3**.

Traditionally, disulfide formation is carried out by air oxidation at neutral or basic pH in aqueous media.[1,2,3,4] This is a very slow method and generally requires a high dilution of the peptide or protein to be effective. Similarly, the thiol–disulfide interchange reaction, with reduced and oxidised glutathiones, is also time consuming and requires high dilution conditions.[5] Despite these drawbacks the long reaction time results in the production of the thermodynamically preferred confomer, which is advantageous when forming two or more disulfides.

Stronger oxidising agents such as iodine and potassium ferricyanide ($K_3Fe(CN)_6$) can be employed in the formation of a single disulfide.[6,7,8] However, methionine, tyrosine, tryptophan and histidine are susceptible to undergo side reactions under these conditions.[9] As a result, the concentration of oxidising agent used must be kept to a minimum and the reaction carried out cautiously.

Tam[10] has detailed a procedure that can be used over a wide pH range and at high concentration to effect disulfide formation. Dimethylsulfoxide (DMSO) is used as a mild oxidant generating water and dimethylsulfide as harmless, easily separable, by-products. An optimum concentration of 20% DMSO by volume, in a variety of solutions, effects complete disulfide formation in peptides in 0.5 to 4 hours. Studies have shown that only methionine is prone to undergo oxidation, sulfoxide formation being observed when working at a pH of less than 3.

2.3 Current Cysteine Protecting Groups

The soft nucleophilic nature of the thiol group facilitates selective *S* protection by direct alkylation of cysteine. Several protecting groups have been proposed for cysteine. **Table 2.1** details those which have been investigated for use in conjunction with Fmoc methodology.

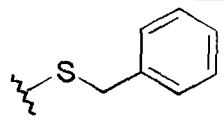
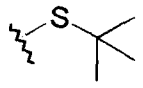
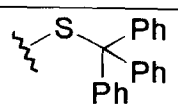
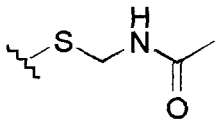
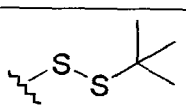
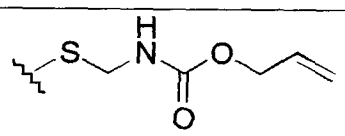
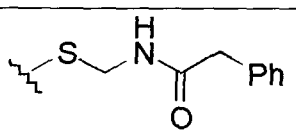
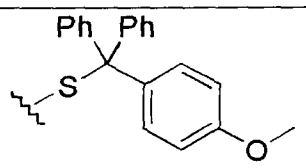
Protecting Group	Cleavage Conditions
 Benzyl[11,12]	Sodium in liquid ammonia.
 *t*-Butyl[13]	Thallium (III) acetate,[14] causes side reaction with Trp. Mercuric acetate,[15] causes Met oxidation to sulfoxide. Both reagents present toxicity problems.
 Trityl[16,17]	Complete cleavage with 95% v/v TFA/H_2O Deprotection with iodine produces the disulfide in one step.[6] Possible side reaction with Trp residues.[9]
 Acetamidomethyl (Acm)[18]	Mercuric acetate at pH 4. Silver trifluoromethane sulfonate (AgOTf),[19] thallium (III) acetate,[14] iodine[20] and TFMSA-DMSO-TFA[21] have all been reported.
 *t*-Butyl sulfenyl[22]	Thiols and phosphines effect complete cleavage before or after cleavage of the peptide from the resin.
 Allyloxycarbonylaminomethyl (Allocam)[23]	$PdCl_2(PPh_3)_2$ mediated cleavage with Bu_3SnH and AcOH. AcOH required to prevent formation of allyl thio ether. Partial cleavage is observed in TFA.
 *S*-Phenylacetamidomethyl (Phacm)[24]	Enzymatic cleavage with *Penicillin Amidohydrolase* at neutral pH.
 Methoxy trityl[25]	3% TFA in DCM-Triethylsilane (95 – 5) in 30 mins. Cleavage can be accomplished with the peptide still bound to the resin.

Table 2.1 Fmoc Compatible Cysteine Protecting Groups

2.4 Problems Encountered with Cysteine Residues in SPPS

Perhaps the most serious problem encountered with cysteine residues is racemisation, due to the acidity of the α-proton. It has been reported that Fmoc-Cys(Trt)-OH may undergo considerable levels of racemisation when 2-(1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluroniumtetrafluoroborate (TBTU), HOBt and DIEA mediated activation and coupling is employed.[26] Coupling *via* the neutral pre-formed symmetrical anhydride yielded material with no detectable level of epimerization. Thus, it can be concluded that the epimerization occurs *via* a base induced mechanism. Extensive racemisation studies have been carried out using the superior coupling reagent HOCt.[27] No racemisation of Fmoc-Cys(Trt)-OH has been observed under these conditions.

Peptides containing *C*-terminal cysteine are also prone to undergo an undesired side reaction when treated with piperidine.[28] During routine cleavage of the Fmoc group, the piperidine may also attack the α-proton of the *C*-terminal cysteine residue. This results in elimination generating a dehydroalanine derivative (**2.1**). A second molecule of piperidine then attacks the alkene, producing a piperidine adduct (**2.2**).

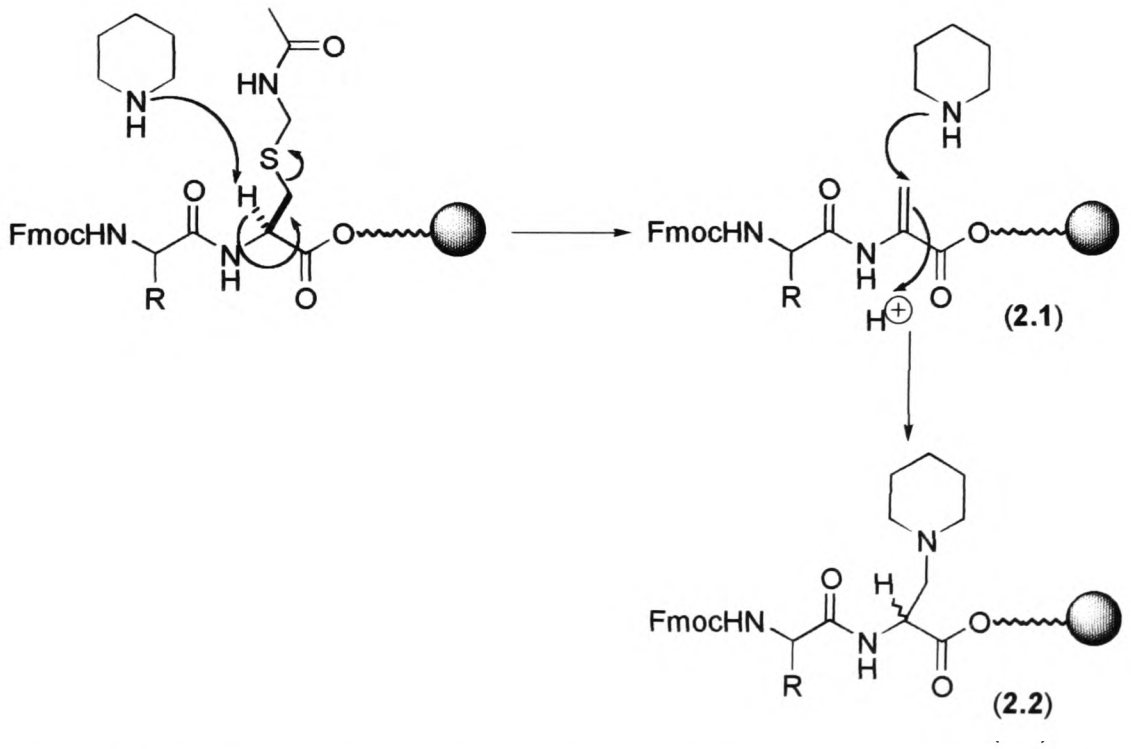


Figure 2.1 Piperidine Attack of C-Terminal Cysteine Residues

Attempts were made to eliminate this reaction using the bulky non-nucleophilic base 1,8-diazobicyclo[5.4.0]undec-7-ene (DBU). However, the dehydroalanine derivative was still isolated from the reaction. No method has yet been developed to completely eliminate this undesired reaction. Residues protected with either the trityl or Acm group were found to be susceptible to this reaction, although the problem is more pertinent for Acm protected residues. Only *C*-terminal cysteine residues undergo this reaction – any other cysteine residues in the peptide chain remain untouched.

2.5 Pseudo Prolines

During peptide synthesis formation of secondary structure is often paralleled by a significant decrease in the solubility of the peptide chain. As a result the resin effectively shrinks, leading to poor accessibility of reagents and, an often substantial, decrease in coupling efficiency. Incorporation of a proline residue into a potentially secondary structure forming region disrupts the formation of helices and β-sheets, thus the solubility of the peptide chain and coupling kinetics are improved.[29]

Mutter and co-workers have introduced structure-disrupting agents, known as pseudo prolines (ΨPro), to circumvent this problem in SPPS.[30,31] Condensation of serine, threonine or cysteine residues with an aldehyde or ketone yields an oxazolidine or thiazolidine ring. These exhibit structure disrupting properties comparable to proline residues when incorporated into peptide chains.

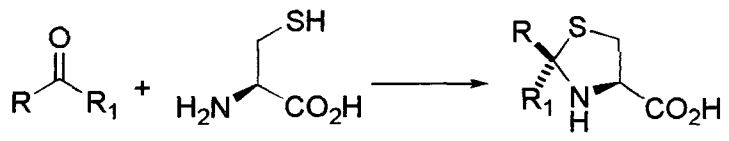


Figure 2.2 Thiazolidine Formation from Condensation of Cysteine and a Ketone

The cyclic system prefers to form a cis-amide bond with the subsequent residue, resulting in a kink in the peptide chain.[30]

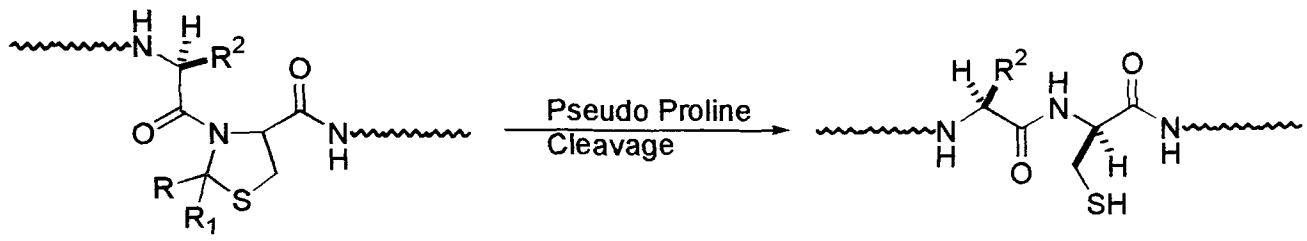


Figure 2.3 Kink Formed in Peptide Chain

By cleaving the oxazolidine or thiazolidine ring at the end of the synthesis, the parent amino acid can be regenerated. Thus, pseudo prolines also serve as a temporary side chain-protecting group for the hydroxyl group of serine and threonine and the thiol group of cysteine.

By varying the nature of the *C*-2 substituents the chemical stability of the ring can be altered.[32] This can be seen in the following cysteine protected derivatives.

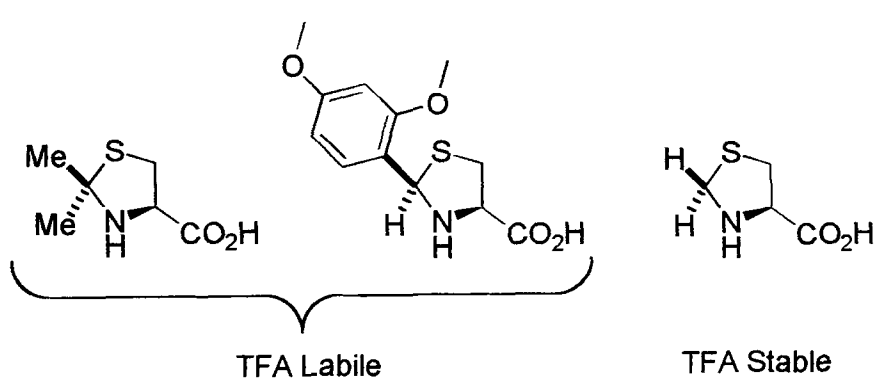


Figure 2.4 Stability of Thiazolidine Derivatives

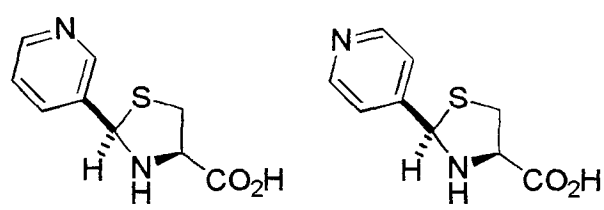
2.6 Outline of Research

It would be desirable to develop an acid and base stable protecting group for the thiol functionality of cysteine residues which could readily be removed in a subsequent chemical step. Thus an orthogonal protecting group to those already described (**Table 2.1**) could be synthesised and employed for the synthesis of peptides and proteins containing two or more disulfides. Ideally, such a protecting group should avoid the use of toxic reagents in cleavage, such as the thallium and mercuric salts required for the cleavage of the Acm and *t*-Bu groups.[14,15,18]

The general requirements for a protecting group can be outlined as follows:

1. Suppress the intrinsic reactivity of the amino acid side chain.
2. Not introduce any new and undesirable reactivity.
3. Be completely stable to the reaction conditions of both acylation and deprotection steps.
4. Ultimately be removable under mild conditions, which do not damage the polypeptide chain.

It was proposed that introduction of a pyridine ring into the pseudo proline system would generate a group stable to both acidic and basic conditions. Cleavage would then be facilitated under reductive conditions. The thiazolidines shown below were both proposed as potential cysteine protecting groups.



(2.3) **(2.4)**

Figure 2.5 Proposed Cysteine Protecting Groups

2.7 Pseudo Proline Derived Cysteine Protecting Groups

2.7.1 Synthesis of the Protecting Groups

Both these compounds were readily synthesised from condensation of *L*-cysteine with the corresponding pyridine carboxaldehyde, using the procedure of Fülöp and co-workers,[33] as outlined in **Figure 2.6**.

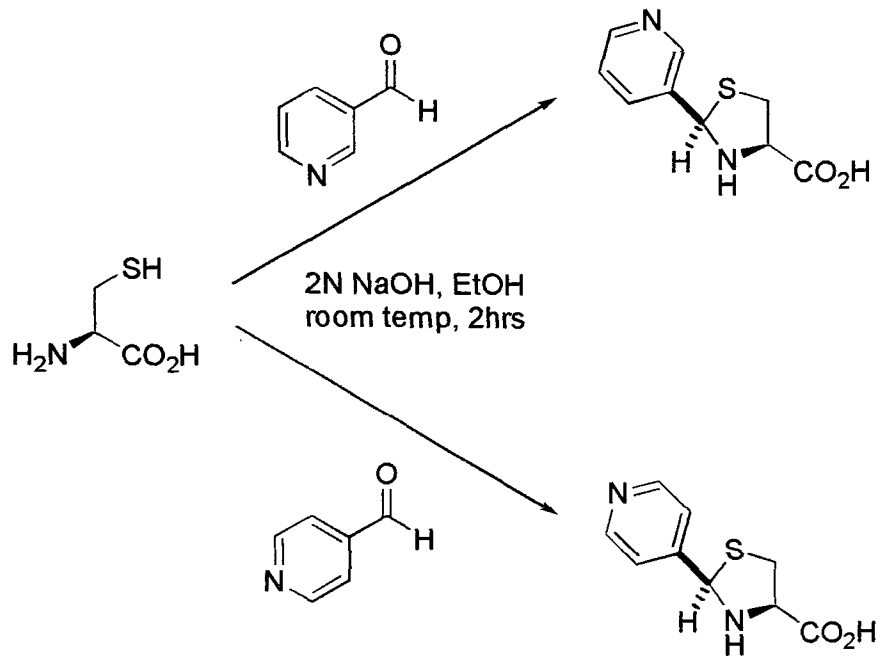


Figure 2.6 Synthesis of the Protecting Groups

Both compounds were obtained as a mixture of diastereomers, due to the introduction of the new chiral centre at *C*-2. Initial attempts to separate these proved unsuccessful since it can be seen that the diastereomers interconvert freely in solution (**Figure 2.7**).[34,35] Although undesirable, the chirality of the amino acid itself is not affected and the protecting group will be removed in a subsequent step.

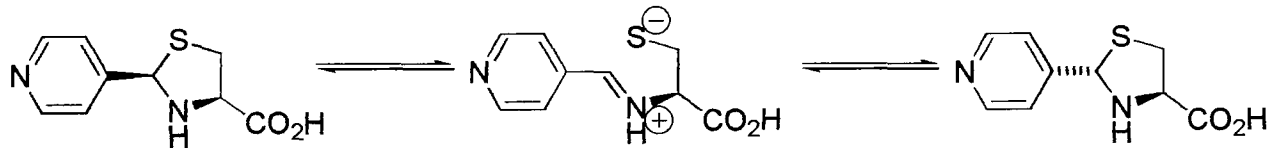


Figure 2.7 Interconversion of Pseudo Prolines in Solution

Coupling of amino acid derivatives to a growing peptide chain containing an *N*-terminal pseudo proline generally results in low yields. This is due to the decreased nucleophilicity of the ring nitrogen of the oxazolidine/thiazolidine.[31] This problem can be overcome by coupling a pre-formed dipeptide of the form Fmoc-Xaa-ΨPro-OH. These dipeptides can readily be incorporated into the growing peptide chain using standard coupling techniques. Racemisation is not found to be a problem due to the proline like structure of the *C*-terminal residue.

2.7.2 Attempted Cleavage of the Protecting Group

2.7.2.1 3-Pyridyl Isomer

Cleavage of the protecting group was attempted using sodium cyanoborohydride ($NaBH_3CN$) in aqueous acetic acid. $NaBH_3CN$ is known to be a selective reducing agent for the production of 1,4-dihydropyridines.[36,37,38] If this intermediate could be generated, then it was thought likely that the protecting group would cleave *via* the mechanism outlined in **Figure 2.8**.

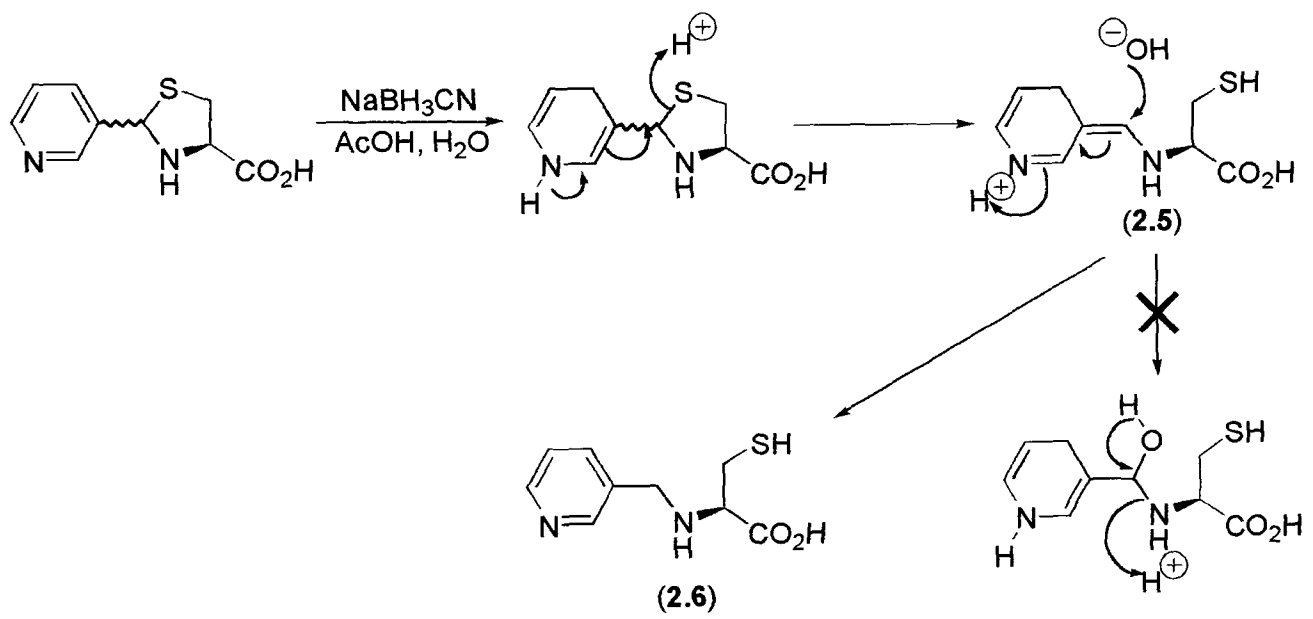


Figure 2.8 Proposed Cleavage of the 3-Pyridine Protecting Group

Analysis of the product produced indicated that only one of the thiazolidine ring bonds had cleaved, namely the *C-S* bond as indicated in **Figure 2.8**. The enamine intermediate (**2.5**) did not undergo attack by the hydroxyl ion as expected, but rather rearranged to form the more stable aromatic system (**2.6**). Consequently cleavage of the *C-N* bond was not observed. Attempts were made to drive this reaction to completion by introducing an excess of reducing agent, however complete cleavage of the system was not observed.

2.7.2.2 4-Pyridyl Isomer

Cleavage of the 4-pyridine isomer was attempted using activated zinc dust as the reducing agent.[39] This method has successfully been employed for the cleavage of the isonicotinyloxycarbonyl (*i*NOC) group, which can be used for protection of the ε-amino group of lysine residues.[40]

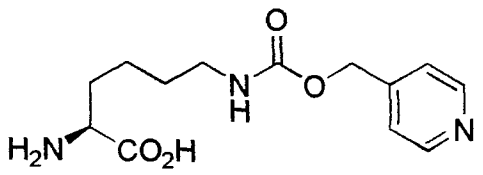


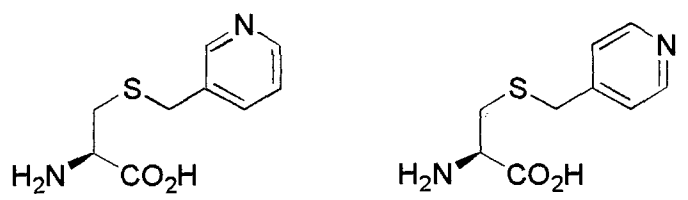
Figure 2.9 Lysine Protected with the Isonicotinyloxycarbonyl Group

However, the product of the reaction was again the partially cleaved material.

The described results show that the pyridine ring does provide an acid and base stable protecting group for the thiol functionality of cysteine residues. The somewhat disappointing results for cleavage of this group do, however, indicate that partial cleavage of the system is possible. Since this implies that reduction of the aromatic system is occurring under these conditions, it should be possible to use these results in the development of a modified protecting group.

2.8 Picolyl Thioethers

Picolyl thioethers of the form outlined below (**Figure 2.10**) should fulfil the requirements for the protecting group. The pyridine ring again providing the acid and base stability. From the results obtained with the thiazolidines reductive conditions should facilitate cleavage of the *C-S* bond.



(2.7) **(2.8)**

Figure 2.10 Picolyl Thioethers as Potential Cysteine Protecting Groups

S-4-Picolyl-*L*-cysteine (**2.8**) has previously been reported as a protected derivative for SPPS by Young and co-workers,[41] who have demonstrated that quantitative cleavage can be accomplished by electrolytic reduction. Despite the obvious advantages that this group appears to offer there is very little use of it reported in the literature.

A further advantage of these picolyl thioethers is that the amino group of the cysteine can be protected with the Fmoc group, allowing incorporation of the protected amino acid into peptide chains using conventional coupling procedures.

2.8.1 Synthesis of the Protecting Group

Synthesis of both the 3-picolyl (**2.9**) and 4-picolyl (**2.10**), Fmoc protected, amino acid can be readily accomplished in two steps, as outlined below, **Figure 2.11**.

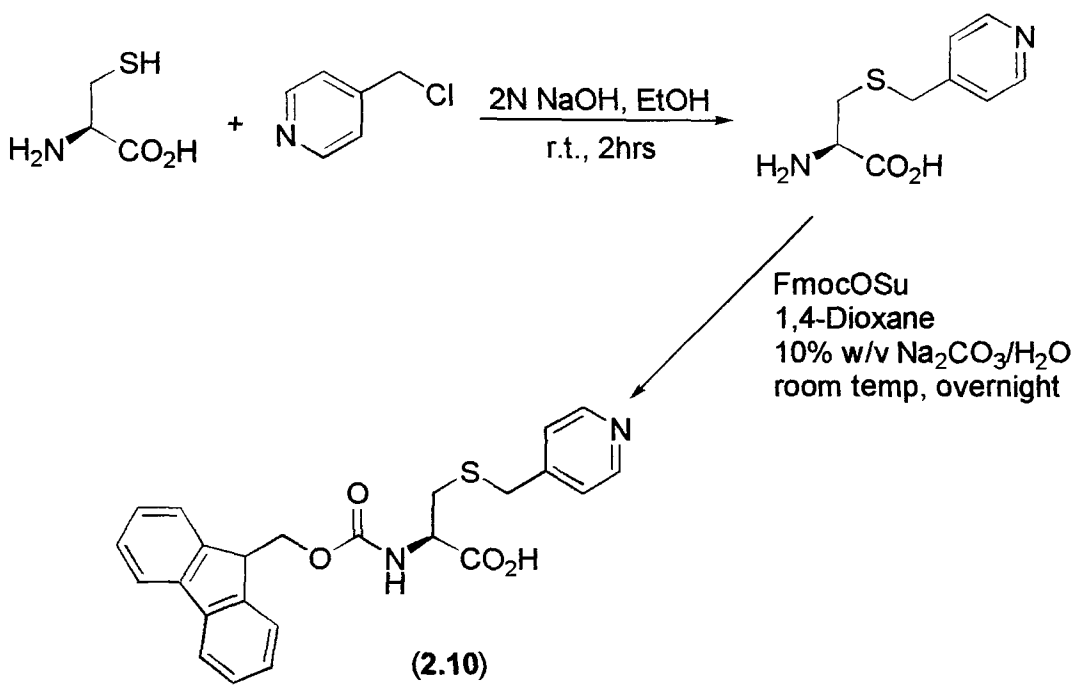


Figure 2.11 Synthesis of Fmoc-Cys(Pic)-OH

2.8.2 Preparation of a Test Peptide

To establish if Fmoc-Cys(Pic)-OH could be utilised in automated SPPS a test peptide was synthesised, namely the *C*-terminus of the vasoconstrictor Endothelin-1 (ET-1).

Figure 2.12 Sequence of Test Peptide

Using standard procedures both isomers were found to couple in greater than 99% efficiency. Both peptides were subjected to acidiolytic cleavage and purified by semi-preparative HPLC. Amino acid analysis (AAA) indicated the picolyl group is stable to prolonged acid treatment. The protected amino acid was found to elute

intact from the ion exchange column with retention properties similar to histidine. This provides a useful means of analysis for picolyl containing peptides.

2.8.3 Attempted Cleavage of the Protecting Groups

2.8.3.1 *S*-3-Picolyl-*L*-Cysteine (2.11)

Cleavage of the 3-picolyl group was attempted using the $NaBH_3CN$ conditions employed for the pseudo proline cleavage. Unfortunately, no cleavage of the protecting group was observed. This was a somewhat surprising result, since partial cleavage of the thiazolidine was observed under identical conditions.

Electrochemical studies on *S*-3-picolyl-*L*-cysteine indicate that the reduction of the pyridine ring does indeed take place. However, the reduction is rapidly reversible, therefore, it appears likely that the 1,4-dihydropyridine is formed, but spontaneously re-aromatises before cleavage of the system can occur.

2.8.3.2 *S*-4-Picolyl-*L*-Cysteine (2.12)

Cleavage was attempted using the established conditions of zinc dust in aqueous acetic acid. The reaction was monitored by HPLC and it can be seen that the peak corresponding to the protected peptide, with retention time 16.4 minutes, started to disappear and a new peak appeared, with retention time 17.8 minutes, indicating that cleavage of the protecting group was successful.

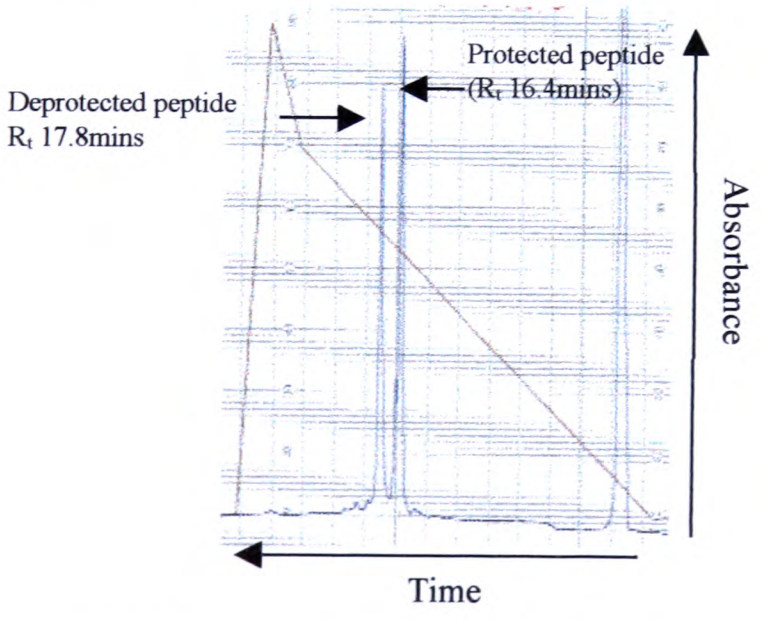


Figure 2.13 HPLC Trace Showing Cleavage of Picolyl Group

A mass of 1210.36 was observed by MALDI-ToF mass spectrometry, corresponding to the deprotected peptide. A positive Ellmans' Test, which detects free thiols, confirmed that cleavage of the picolyl group had indeed occurred.

Electrochemical studies on this material indicated that the reduced form of this isomer was sufficiently stable to allow the cleavage to proceed to completion.

2.8.4 Racemisation Study

The 4-picolyl group fulfilled the initial requirements for a cysteine protecting group. However, as mentioned previously, racemisation is an undesired side reaction observed for a number of protected cysteine derivatives. In order to establish if racemisation occurred, using this protected cysteine derivative, the tri-peptide H_2N-Ala-Cys(Pic)-Gly-OH (**2.13**) was synthesised using HOCt mediated coupling. Racemisation of the cysteine residue can be identified by ^1H NMR spectroscopy, since the characteristic doublet of the methyl group of the alanine residue will resonate as two characteristic doublets at ~1.3ppm.[27]

As can be seen from the ^1H NMR spectrum (**Figure 2.14**), there is no detectable level of racemisation of the cysteine residue. A sample of the peptide was prepared with the D-Cys(Pic) residue and used to spike the ^1H NMR sample. This confirmed that it was possible to identify diastereomers using this technique.

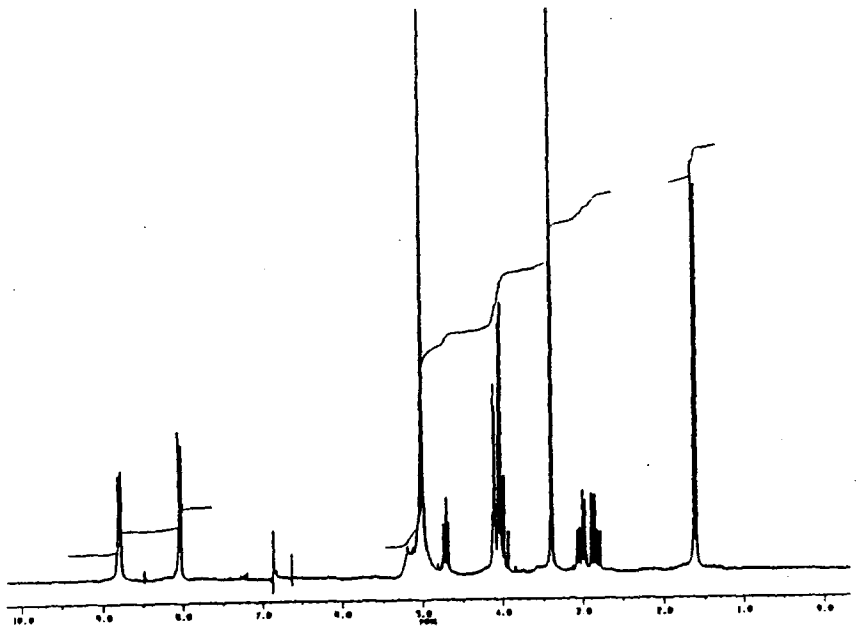


Figure 2.14 ^1H NMR Spectrum of Ala-Cys(Pic)-Gly

2.9 Application to Synthesis

2.9.1 Oxytocin

In order to demonstrate the synthetic utility of Fmoc-Cys(Pic)-OH in SPPS the synthesis of oxytocin was undertaken. Oxytocin is a nine residue peptide containing two cysteine residues. This would enable oxidation to be attempted, to ensure that the picolyl group does not introduce any features, which would later impair disulfide bond formation.

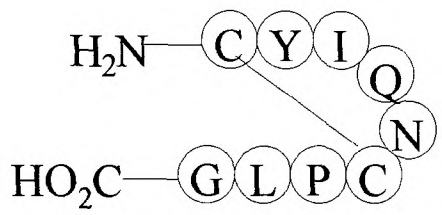


Figure 2.15 Linear Sequence of Oxytocin

Oxytocin was synthesised using standard SPPS techniques, all amino acids were single coupled as the active HOCt ester. No drop in coupling efficiency was seen at either cysteine residue, as demonstrated in the deprotection profile below.

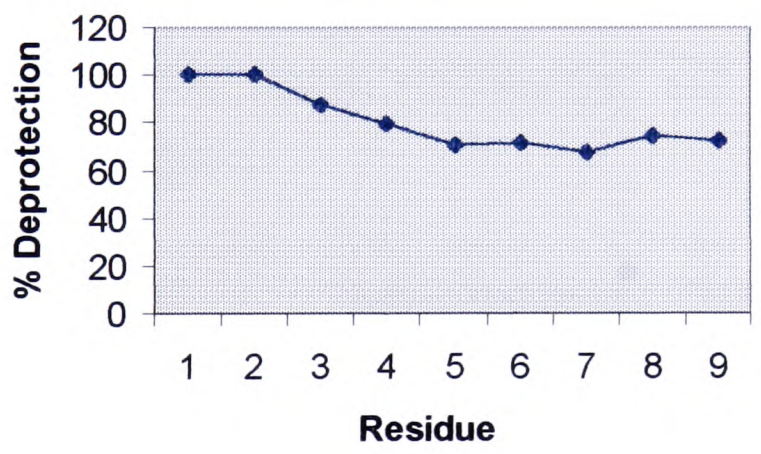


Figure 2.16 Fmoc Deprotection Profile of Oxytocin

Cleavage of the picolyl protecting groups was achieved using zinc dust, complete cleavage observed in 30 minutes as shown in the HPLC trace of the reaction mixture (**Figure 2.17**).

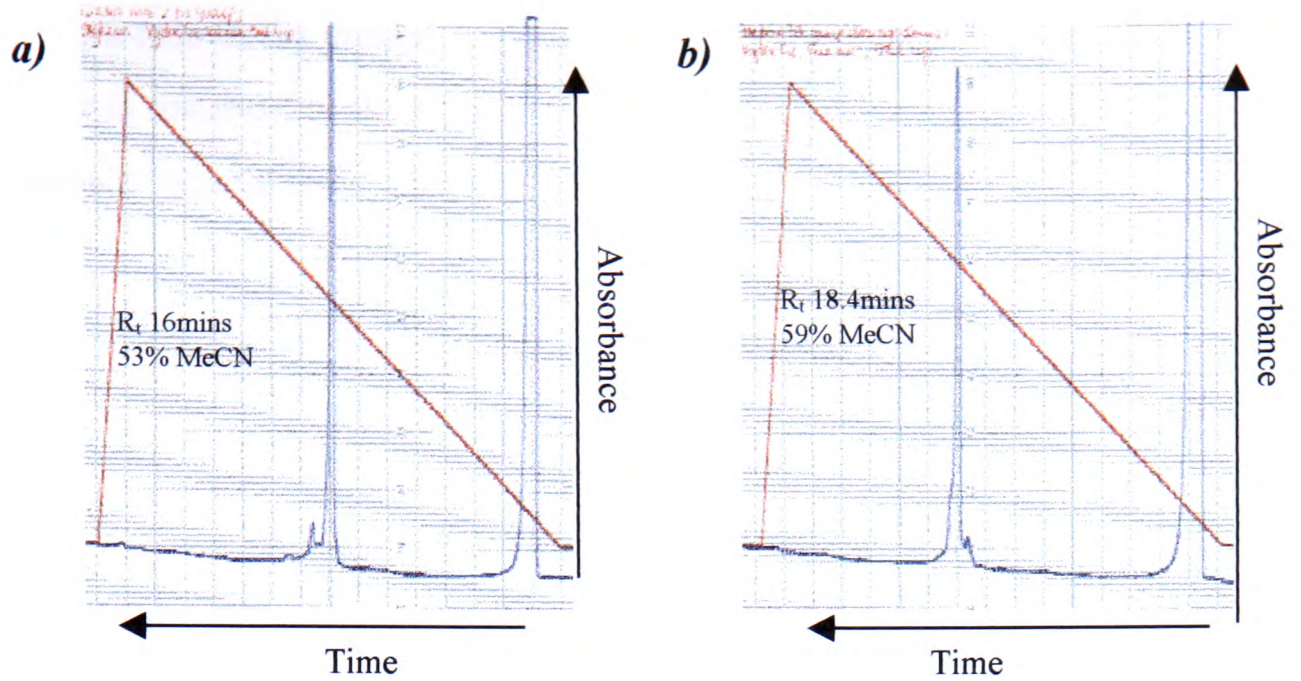


Figure 2.17 HPLC of Oxytocin a) Before and b) After Picolyl Group Cleavage

Disulfide formation was attempted using 20% v/v DMSO in TFA. Monitoring of the reaction by HPLC gave a good indication of the progress of the disulfide formation. After 4 hours the *S-S* formation was deemed complete and the folded material was isolated by semi-preparative HPLC.

All products were analysed by amino acid analysis and electrospray mass spectroscopy (ES-MS), **Table 2.2**.

OXYTOCIN	**ES-MS**	**AMINO ACID ANALYSIS**
Picolyl Protected (**2.14**)	1192.7 (M^+) **MW** 1192.41	Asp_1 (1.03), Glu_1 (1.08), Gly_1 (1.03), Ile_1 (0.97) Leu_1 (1.04), Tyr_1 (0.83), Pro_1 (1.03)
Deprotected Material (**2.15**)	1010.8 (M^+) **MW** 1010.19	Asp_1 (0.87), Glu_1 (1.07), Gly_1 (1.28), Cys_2 (0.86), Ile_1 (1.03), Leu_1 (1.39), Tyr_1 (1.08), Pro_1 (1.53)
Folded Material (**2.16**)	1008.6 (M^+) **MW** 1008.17	Asp_1 (0.97), Glu_1 (1.04), Gly_1 (1.01), Cys_2 (0.35), Ile_1 (0.96), Leu_1 (1.03), Tyr_1 (0.98), Pro_1 (1.00)

Table 2.2 Results for Oxytocin Synthesis

2.9.2 Endothelin-1

The endothelins are a family of peptides, which have potent and well characterised effects on the cardiovascular system.[42] Of these endothelin-1 (ET-1) is the major isoform, derived from the precursor Big ET-1. It is thought that the endothelins may

also have a role in the central nervous system acting as neurotransmitters.[43] ET-1 in particular is involved in the regulation of conditions associated with vasoconstriction, for example hypertension, chronic heart failure and renal failure.[44,45]

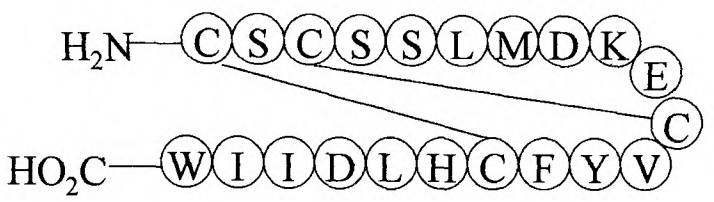


Figure 2.18 Linear Sequence of ET-1

Endothelin-1 (ET-1) is a 21 residue peptide containing two disulfide bridges. Therefore, it provided an interesting target to study the orthogonality of the picolyl group in conjunction with the trityl group. The trityl group is acid labile, cleaving during treatment of the peptide-resin with TFA. Thus it should be possible to form the disulfide from the two free thiol groups generated at this stage. Cleavage of the picolyl groups will then liberate the remaining thiols, which can then be oxidised to produce the unambiguously folded material.

From the deprotection profile it can be seen that the synthesis appeared to suffer a substantial drop, however despite repeating the synthesis with double coupling cycles this was not improved.

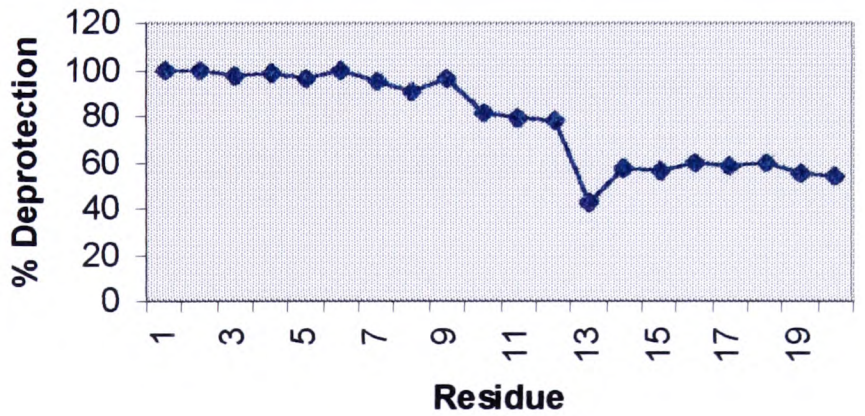


Figure 2.19 Deprotection Profile of ET-1 with 2 Picolyl Groups

Purification of the material was achieved with the introduction of the tetrabenzo[*a,c,g,i*]fluorenyl-17-methoxycarbonyl (Tbfmoc) moiety at the *N*-terminus of the peptide. This group has been developed to facilitate separation of truncated sequences from the desired material.[46] By exploiting the high affinity of this planar

aromatic moiety for charcoal, sequences containing the Tbfmoc group can be adsorbed onto the charcoal.[47,48] Material labelled with the Tbfmoc group is readily identified since the group has a characteristic absorbance at 364nm. Truncates and any other non-labelled impurities can be washed away leaving the pure peptide bound to the charcoal. Cleavage of the Tbfmoc group from the peptide can then be achieved on treatment with piperidine, liberating the purified peptide.

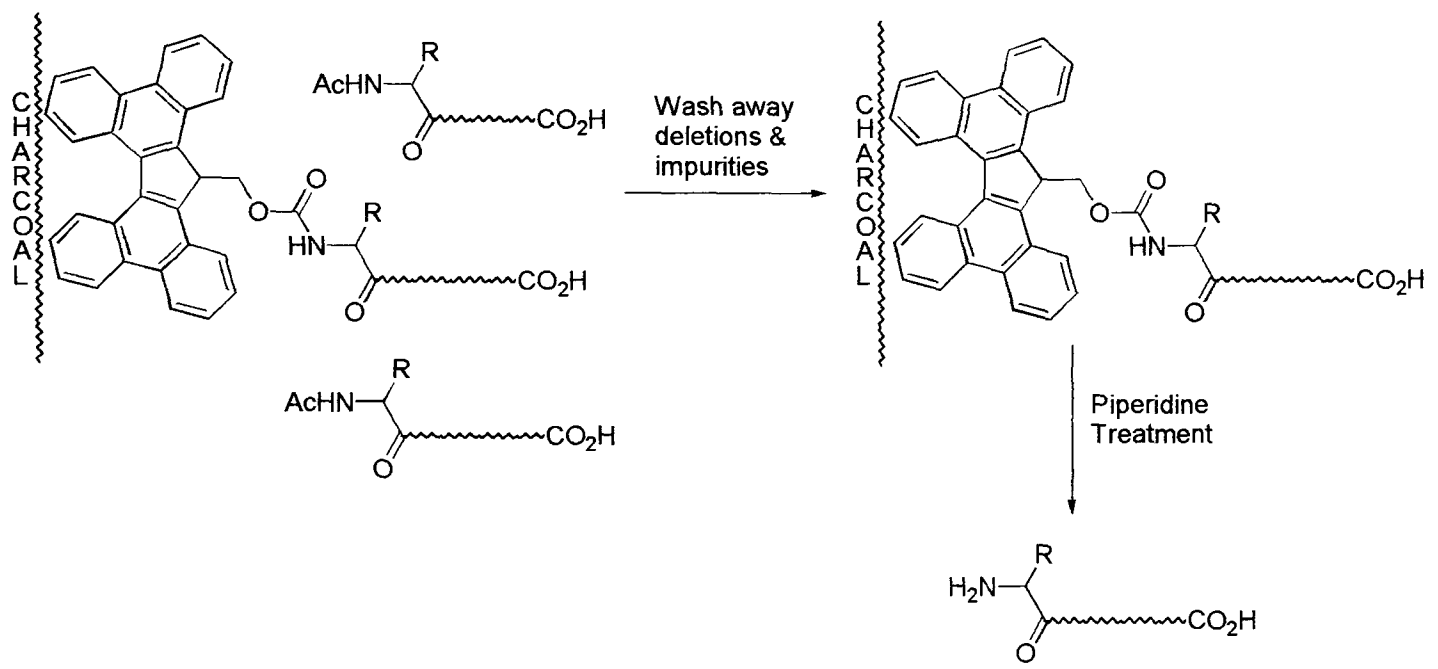


Figure 2.20 General Outline of Tbfmoc Affinity Purification

Following purification of ET-1, disulfide formation was effected using DMSO mediated oxidation. However, upon subsequent treatment of the peptide with zinc dust to effect picolyl group cleavage, it appeared that the disulfide bond already formed was also undergoing cleavage. This perhaps should not have been totally unexpected, considering the reducing conditions used to cleave the picolyl group.

A sample of commercially available *L*-cystine was subjected to the picolyl group cleavage conditions. Analysis of the product produced from the reaction proved that the zinc dust was indeed a strong enough reducing agent to effect cleavage of the disulfide bond.

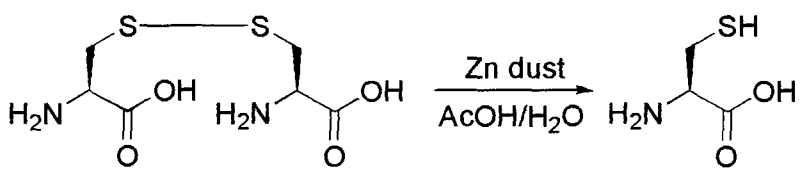


Figure 2.22 Reduction of L-Cystine

Although this was a disappointing result, it did prove that the picolyl group is compatible with the Tbfmoc affinity purification method for synthetic peptides. In order to demonstrate this method fully the synthesis of ET-1 was repeated, this time employing picolyl protection for all four cysteine residues. The material was purified by Tbfmoc / charcoal affinity purification and the picolyl groups cleaved in the manner outlined above. Deprotection appeared to be complete after one hour.

Folding of the material was then achieved using DMSO mediated oxidation. Although this could generate three possible products, due to mixed disulfide formation, only two products were detected by HPLC. The major component identified as the correctly oxidised material *via* co-injection with a commercial standard.

ENDOTHELIN	**ES-MS**	**AMINO ACID ANALYSIS**
Picolyl Protected (**2.17**)	2859.9 (M^+) **MW** 2859.89	Asp_2 (2.04), Ser_3 (2.20), Glu_1 (1.25), Val_1 (1.25), Met_1 (0.78), Ile_2 (1.88), Leu_2 (2.20), Tyr_1 (1.10), Phe_1 (1.10), His_1 (1.10), Lys_1 (0.94)
Deprotected Material (**2.18**)	2495.1 (M^+) **MW** 2495.89	Asp_2 (1.77), Ser_3 (1.27), Glu_1 (1.18), Cys_4 (0.52), Val_1 (1.18), Met_1 (0.69), Ile_2 (2.13), Leu_2 (2.10), Tyr_1 (0.69), Phe_1 (1.12), His_1 (1.02), Lys_1 (1.21)
Folded Material (**2.19**)	2491.1 (M^+) **MW** 2491.89	Asp_2 (2.02), Ser_3 (1.96), Glu_1 (1.19), Cys_4 (1.29), Val_1 (1.04), Met_1 (0.97), Ile_2 (1.91), Leu_2 (2.07), Tyr_1 (0.64), Phe_1 (0.97), His_1 (0.98), Lys_1 (0.96)

Table 2.3 Results for Endothelin-1 Synthesis

2.10 Conclusions

The picolyl group is an effective acid and base stable protecting group for the thiol functionality of cysteine residues. Incorporation of picolyl protected cysteine residues into growing peptide chains can be readily achieved utilising conventional SPPS techniques, HOCt mediated coupling causing no detectable levels of racemisation. Unfortunately the reducing conditions required to cleave the picolyl

group do cleave any disulfide bonds already present in the peptide. As a result, the picolyl group cannot be employed as an orthogonal protecting group in conjunction with the trityl group. However, it is possible that orthogonality could be achieved in conjunction with non-acid labile cysteine protecting groups, for example Acm. Obviously adopting this approach would necessitate cleavage of the picolyl group first. The excess zinc dust can be removed from the peptide, following deprotection, simply by filtration. The zinc salts formed during the reaction can be separated from the peptide by semi-preparative HPLC or gel filtration. Thus, harmful by-products, often accumulated during cleavage of other protecting groups, are avoided.

Additionally, the solubilising effect the picolyl group confers on the peptide chain aids in purification procedures. This, coupled with the stability of the picolyl group, means that the picolyl group has ideal properties for use in fragment coupling reactions using minimally protected peptide fragments. This area will be detailed later in **Chapter 4**.

1. C. H. Li, D. Yamashiro, D. Gospodarowicz, S. Kaplan, G. Van Vilet, *Proc. Natl. Acad. Sci. U.S.A.*, 1983, **80**, 2216.

2. H. Yajima and N. Fujii, *J. Am. Chem. Soc.*, 1981, **103**, 5867.

3. J. P. Tam, *Int. J. Pept. Protein Res.*, 1987, **21**, 421.

4. J. P. Tam, M. A. Sheikh, D. S. Salomon and L. Ossowski, *Proc. Natl. Acad. Sci. U.S.A.*, 1986, **83**, 8082.

5. A. K. Ahmed, S. W. Schaffer and D. B. Wetlaufer, *J. Biol. Chem.*, 1975, **250**, 8477.

6. B. Kamber, A. Hartmann, K. Eisler, B. Riniker, H. Rink, P. Sieber and W. Rittel, *Helv. Chim. Acta.*, 1980, **63**, 899.

7. S. Veber, J. D. Milkowski, S. L. Varga, R. G. Denkewalter and R. Hirschmann, *J. Am. Chem. Soc.*, 1972, **94**, 5456.

8. D. B. Hope, S. V. V. Murti and V. Du Vigneaud, *J. Biol. Chem.*, 1962, **237**, 1563.

9. P. Sieber, B. Kamber, B. Riniker and W. Rittel, *Helv. Chim. Acta.*, 1980, **63**, 2358.

10. J. P. Tam, C. Wu, W. Liu and J. Zhang, *J. Am. Chem. Soc.*, 1991, **113**, 6657.

11. R. H. Sifferd and V. du Vigneaud, *J. Biol. Chem.*, 1935, **108**, 753.

12. J. L. Wood and V. du Vigneaud, *J. Biol. Chem.*, 1939, **130**, 109.

13. F. M. Callahan, G. W. Anderson, R. Paul and J. E. Zimmerman, *J. Am Chem. Soc.*, 1963, **85**, 201.

14. N. Fujii, A. Otaka, S. Funakoshi, K. Bessho and H. Yajima, *J. Chem. Soc., Chem. Commun.*, 1987, 163.

15. S. M^{C}Curdy, *Peptide Research*, 1989, **2**, 147.

16. G. Amard, R. Heymès and L. Velluz, *Bull. Soc. Chim.*, 1956, 698.

17. L. Zervas and D. M. Theodoropoulos, *J. Am. Chem. Soc.*, 1956, **78**, 1359.

18. D. F. Veber, J. D. Milkowski, S. L. Varga, R. G. Denkewalter and R. Hirschmann, *J. Am. Chem. Soc.*, 1972, **94**, 5456.

19. N. Fujii, A. Otaka, T. Watanabe, A. Okamachi, H. Tamarmura, H. Yajima, Y. Inagaki, M. Nomizu and K. Asano, *J. Chem. Soc., Chem. Commun.*, 1989, 283.

20. B. Kamber, *Helv. Chim. Acta.*, 1971, **54**, 927.

21. H. Tamamura, T. Ishihara, A. Otaka, T. Koide, K. Miyoshi, T. Ibuka and N. Fujii, *J. Chem. Soc., Perkin Trans. I.*, 1996, 1911.

22. R. Eritja, J. P. Ziehler-Martin, P. A. Walker, T. D. Lee, K. Legesse, F. Albericio and B. E. Kaplan, *Tetrahedron*, 1987, **43**, 2675.

23. A. M. Kimborigiula, A. Merzouk, F. Guibé and A. Loffet, *Tetrahedron Lett.*, 1994, **35**, 9035.

24. M. Royo, J. Alsina, E. Giralt, U. Slomcyznska and F. Albericio, *J. Chem. Soc., Perkin Trans. I*, 1995, 1095.

25. K. Barlos, D. Gatos, D. Hatzi, N. Koch and S. Koutsogianni, *Int. J. Peptide Protein Res.*, 1996, **47**, 148.

26. T. Kaiser, G. J. Nicholson, H. J. Kohlbau and W. Voelter, *Tetrahedron Lett.*, 1996, **37**, 1187.

27. N. Robertson, L. Jiang and R. Ramage, *Tetrahedron*, 1999, **55**, 2713.

28. J. Lukszo, D. Patterson, F. Albericio and S. A. Kates, *Lett. Peptide Sci.*, 1996, **3**, 157.

29. T. Haack and M. Mutter, *Tetrahedron Lett.*, 1992, **33**, 1589.

30. A. Nefzi, K. Schenk and M. Mutter, *Protein and Peptide Letters*, 1994, **1**, 66.

31. M. Mutter, A. Nefzi, T. Sato, X. Sun, F. Wahl and T. Wöhr, *Peptide Res.*, 1995, **8**, 145.

32. T. Wöhr, F. Wahl, A. Nefzi, B. Rohwedder, T. Sato, X. Sun and M. Mutter, *J. Am. Chem. Soc.*, 1996, **118**, 9218.

33. F. Fülöp, J. Mattinen and K. Pihlja, *Tetrahedron,* 1990, **46**, 6545.

34. G. S. Sheppard, D. Pireh, G. M. Carrera Jr., M. G. Bures, H. R. Heyman, D. H. Steinman, S. K. Davidsen, J. G. Phillips, D. E. Guinn, P. D. May, R. G. Conway, D. A. Rhein, W. C. Calhoun, D. H Albert, T. J. Majoc, G. W. Carter and J. B. Summers, *J. Med. Chem.*, 1994, **37**, 2011.

35. R. Faggiani, H. E. Howard-Lock, C. J. L. Lock and R. Orgias, *Can. J. Chem.*, 1991, **69**, 1.

36. U. Eisner and J. Kuthan, *Chem. Rev.*, 1972, **72**, 1.

37. E. Booker and U. Eisner, *J. Chem. Soc., Perkin Trans. I.*, 1975, 929.

38. D. M. Stout and A. I. Meyers, *Chem. Rev.*, 1982, **82**, 223.

39. K. Tsuda and E. Ohki, *J. Org. Chem.*, 1963, **28**, 783.

40. D. F. Veber, W. J. Paleveda, Y. C. Lee and R. Hirschmann, *J. Org. Chem.*, 1977, **42**, 3286.

41. A. Gosden, R. Macrae and G. T. Young, *J. Chem. Res.(M).*, 1977, 0310.

42. M. Yanagisawa, H. Kurihara, S. Kimura, Y. Tomobe, M. Kobayashi, Y. Mitsui, Y. Yazaki, K. Goto and T. Masaki, *Nature*, 1988, **332**, 411.

43. B. Zhu and J. Herbert, *Neuroscience*, 1996, **71**, 1049.

44. W. G. Haynes and D. J. Webb, *Clin. Sci.*, 1993, **84**, 485.

45. D. J. Webb and W. G. Haynes, *The Lancet*, 1993, **342**, 1439.

46. R. Ramage and G. Raphy, *Tetrahedron Lett.*, 1992, **33**, 385.

47. R. Ramage, S. L. Irving and A. Brown, *Tetrahedron Lett.*, 1993, **34**, 7129.

48. R. Ramage, H. R. Swenson and K. T. Shaw, *Tetrahedron Lett.*, 1998, **39**, 8715.

CHAPTER 3

SPPS of Deglycosylated Human Erythropoieitin

3.1 Background

Erythropoietin (EPO)[1] is an acidic glycoprotein produced primarily in the kidney of adults and the liver of the foetus. It is a member of the extensive cytokine family of proteins,[2] which control a wide range of functions in cells of the lymphoid, hemopoietic and reticuloendothelial systems. EPO is found in low concentration in the plasma and is responsible for the differentiation of erythroid progenitor cells, found in the bone marrow, into mature red blood cells.[4] Red blood cells are required for transport of oxygen around the body to various tissues. In cases of hypoxia, EPO production, and in turn red blood cell production is increased. Conversely, in cases of hyperoxia, a decrease in both EPO and red blood cell production is noted.

The possibility that erythropoiesis was controlled by a hormone was first postulated in 1906 by Carnot and Deflandre,[1] although it was 50 years before substantial evidence was obtained in support of this. Erslev[3] noted an increase in blood reticulocytes and hematocrit levels (% red blood cells in blood volume) when healthy rabbits were injected with large amounts of plasma from anaemic rabbits.

Isolation of EPO proved to be difficult due to the low concentration of the hormone present in the plasma. In 1977 Miyake[4] successfully managed to purify a small quantity of EPO (uhEPO) to homogeneity from the urine of patients with aplastic anaemia. With the advent of modern molecular biology techniques sequencing of the protein enabled both Jacobs[5] and Lin[6] to clone the EPO gene. Expression of this gene in either African green monkey kidney cells (COS-1)[5] or Chinese hamster ovary cells (CHO)[6] produced recombinant human erythropoietin (rhEPO). Subsequent studies have shown rhEPO and uhEPO to be identical.[7]

Kidney failure can cause a decrease in EPO production resulting in anaemia. Treatment of renal failure with blood transfusions can offer temporary relief however, transfusions can cause antibodies to develop which can preclude patients from later receiving a successful kidney transplant.

Amgen successfully initiated clinical trials of rhEPO in 1985, showing that therapeutic doses stimulated consistent increases in patients' hematocrit levels, substantially eliminating the requirement for blood transfusions for dialysis patients. rhEPO has been clinically licensed as EPOGEN® since 1989[8] and most patients undergoing dialysis receive this as part of their treatment. Since its introduction, rhEPO has found application in the treatment of anaemia associated with chronic disease, such as rheumatoid arthritis, tuberculosis, cancer and systematic fungal infections.[9,10] OrthoBiotech also distribute rhEPO under the trade name PROCRIT® for use in conjunction with Zidovudine (AZT) therapy for the treatment of HIV related anaemia.

3.2 Structural Features

The EPO gene encodes a 193 residue pro-protein, comprising of a 27-residue hydrophobic signal sequence, which is cleaved during secretion, and the mature protein consisting of 166 residues (**Figure 3.1**).

Ala Pro Pro Arg Leu Ile Cys Asp Ser Arg Val Leu Glu Arg Tyr Leu Leu
Glu Ala Lys Glu Ala Glu Asn Ile Thr Thr Gly Cys Ala Glu His Cys Ser
Leu Asn Glu Asn Ile Thr Val Pro Asp Thr Lys Val Asn Phe Tyr Ala Trp
Lys Arg Met Glu Val Gly Gln Gln Ala Val Glu Val Trp Gln Gly Leu Ala
Leu Leu Ser Glu Ala Val Leu Arg Gly Gln Ala Leu Leu Val Asn Ser Ser
Gln Pro Trp Glu Pro Leu Gln Leu His Val Asp Lys Ala Val Ser Gly Leu
Arg Ser Leu Thr Thr Leu Leu Arg Ala Leu Gly Ala Gln Lys Glu Ala Ile
Ser Pro Pro Asp Ala Ala Ser Ala Ala Pro Leu Arg Thr Ile Thr Ala Asp
Thr Phe Arg Lys Leu Phe Arg Val Tyr Ser Asn Phe Leu Arg Gly Lys Leu
Lys Leu Tyr Thr Gly Glu Ala Cys Arg Thr Gly Asp Arg

Figure 3.1 Sequence of dhEPO (1-166)

The physiologically active form of EPO circulating in the plasma is actually des-Arg^{166} and it appears that this *C*-terminal processing occurs either intracellularly prior to secretion, or during circulation in plasma.[11] EPO contains two intramolecular disulfide bonds between Cys^{29}-Cys^{33} and Cys^{7}-Cys^{161}. The small disulfide loop between Cys^{29}-Cys^{33} is not essential for secretion,[12] although there is a marked decrease in activity in its absence, indicating that it appears to be essential for correct functioning if the protein.[13]

There are three *N*-linked sugars at Asp$^{24, 38, 83}$ and also one *O*-linked sugar at Ser^{126}. Glycosylation is not essential for *in vitro* activity, although it is required for *in vivo* activity. The endoplasmic reticulum lumen is an oxidising environment that promotes disulfide formation. Here glycosylation may serve to prevent aggregation of the protein and prevent enzymatic degradation from the numerous proteases.[14] EPO only possesses full biological activity *in vivo* when it is sufficiently sialylated,[15] indeed the sialic acid residues account for 40% of the carbohydrate mass.[16] Removal of the terminal sialic acid residues from the carbohydrate promotes metabolic clearance of the hormone, *via* binding of the terminal galactose residues to the hepatic receptors in the liver.[17,18] Overall, carbohydrates are not required for either biological activity or target cell specificity, but serve to prevent premature removal of the hormone from the blood stream.[19]

The tertiary structure of EPO is proposed to be analogous to that of growth hormone (GRH),[20] which has been crystallised and consists of an anti-parallel four α-helical bundle core. It is proposed that EPO consists of four anti-parallel helices with two long and one short loop connections. The areas predicted to be α-helices are highly conserved in sequences of EPO from various mammals.[21] Mutations in these areas lead to secretion of protein with little or no activity,[12] indicating their importance in the structure and function of the protein.

Recently the NMR structure of an EPO mutant has been solved.[22] All glycosylation sites were mutated to lysine to aid the solubility of the protein and MK added at the

N-terminus to increase expression yields in *E. coli*. Activity was shown to be comparable to that of deglycosylated human EPO (dhEPO) expressed in *E. coli*. The structure has been solved as a left handed α-helical bundle with four long helices running in an up-up-down-down direction connected by two long cross over loops and one short loop (**Figure 3.2**). Residues comprising the central core have been shown to be invariant across the different EPO sequences from various species and mutation of any of these core residues has a marked effect on protein folding.

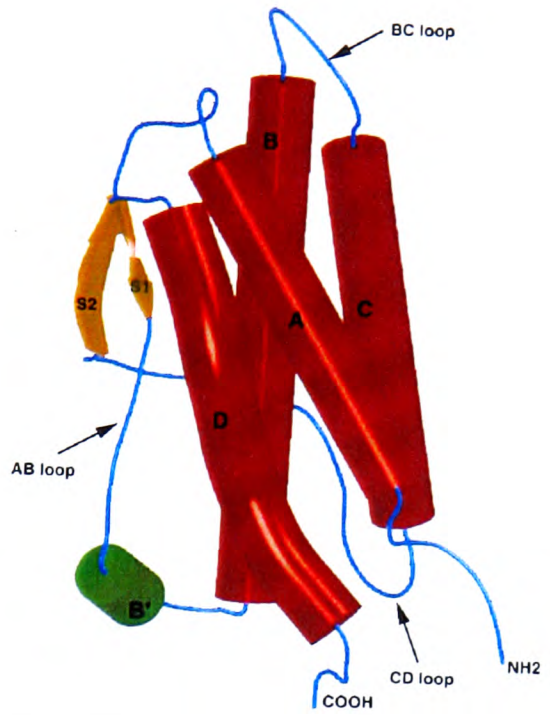


Figure 3.2 NMR Structure of MKLysEPO

3.3 Receptor Binding and Activation

EPO excerts its biological effects through the binding to and orientation of two cell surface EPO receptors (EPObp), which then trigger an intracellular phosphorylation cascade.[23] Cloning and expression of the EPObp gene[24] has shown that the receptor comprises a single polypeptide chain of 508 residues and apparent MW of 55 kDa. EPObp exhibits a number of features common to other members of the cytokine receptor family. It contains a single hydrophobic transmembrane domain bearing four cysteine residues and a WSXWS motif (where X represents any amino acid) close to the transmembrane domain.[25]

The crystal structure of a soluble EPO mutant (Lys substituted at all three *N*-linked glycosylation sites) complexed with the receptor has recently been solved (**Figure 3.3**).[23]

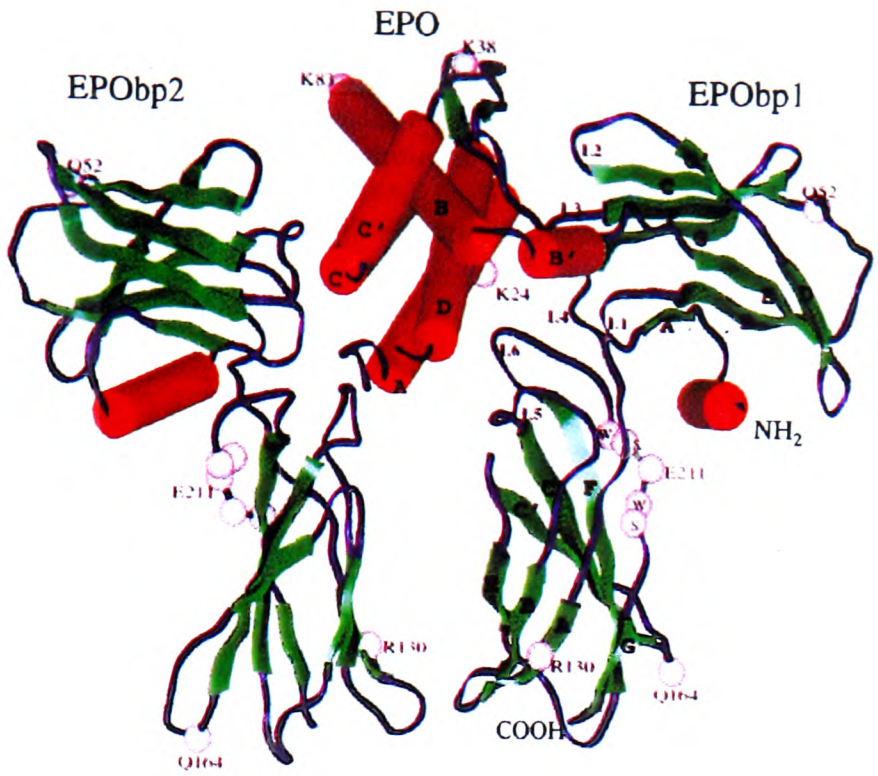


Figure 3.3 MKLysEPO Complexed with Receptor

As predicted from the NMR data,[22] the complex is formed as a 2:1 $EPObp_2$:EPO species, where the two receptors are held together through two regions located on opposite faces of the EPO molecule. These two interfaces are characterised by a high affinity binding at Site 1, and a low affinity binding at Site 2. Site 1 is comprised of residues from helices A, B and D along with part of the AB loop. Site 2 is comprised of residues from helices A and C. Generally the side chain interactions at both binding sites are between the positively charged arginine and lysine residues of EPO and the negatively charged glutamate and aspartate residues of EPObp.

A comparison of the NMR structure of the isolated protein with the $EPObp_2$:EPO complex reveals that the overall topology and alignment of the α-helical bundle is very similar. There are a few differences between the two structures, particularly in the loop regions, however these may be attributed to differences induced upon receptor binding.

3.4 Overview of Research

Due to the advances in methodology for SPPS, outlined in **Chapter 1**, the chemical synthesis of proteins of up to 200 amino acids is now a realistic target. Previous research has demonstrated this is feasible with the synthesis of EPO,[26] interferon-γ (INF-γ)[27] and the catalytic domain of stromelysin (SCD).[28] Unfortunately, the main problem in the chemical synthesis of proteins lies in the subsequent purification steps rather than the actual synthesis. The ultimate aim of chemical protein synthesis is to achieve a sufficient quantity of material for structural analysis, for example x-ray crystallography and NMR spectroscopy. Thus an efficient purification protocol must be devised.

3.5 Chemical Synthesis of dhEPO

Synthesis of dhEPO was carried out on a 0.1mmol scale using Fmoc-Arg(Pmc) functionalised Wang resin. All amino acids were single coupled as the active HOCt ester. Previous work indicated that the coupling efficiency can drop substantially at the last twenty residues of the synthesis,[29] therefore these amino acids were introduced *via* a double couple cycle. The efficiency of the synthesis was monitored by measuring the UV absorbance of the fulvene-piperidine adduct (**Figure 3.4**).[30] The deprotection profile is used as a diagnostic tool to check that there are no major drops in the synthesis, rather than a definitive result of the coupling efficiency.

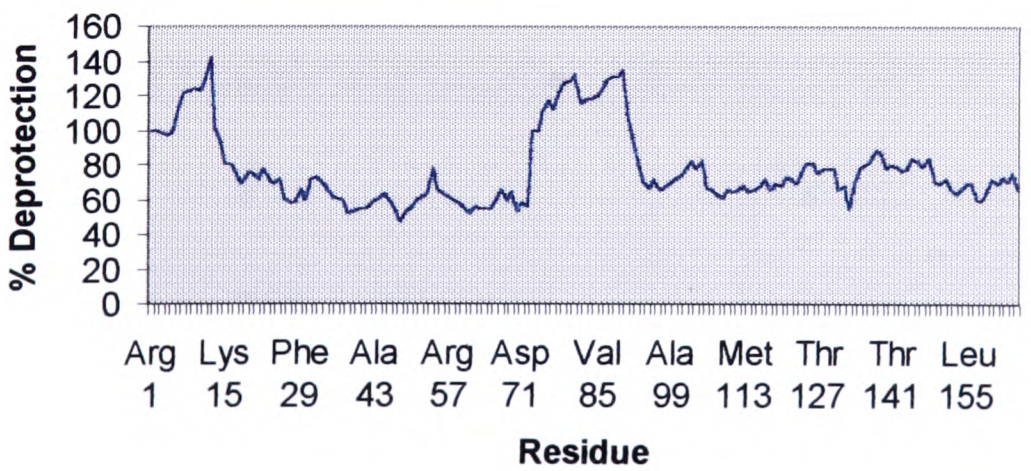


Figure 3.4 Fmoc Deprotection Profile of dhEPO

From the deprotection profile it can be seen that the synthesis appears to have proceeded reasonably well, with no significant drops in coupling efficiency. Regions where the coupling efficiency appears to be greater than 100% are thought to occur due to changes in the swelling properties of the resin, a feature which has been observed in the synthesis of other proteins.

On completion of the synthesis a sample of the resin bound protein was characterised by amino acid analysis. A series of timed reactions were carried out to determine the optimum time for hydrolysis (**Table 3.1**). The results obtained appeared to be in reasonable agreement with the predicted values, suggesting that the synthesis had been successful. Twenty-four hours was chosen as the optimum hydrolysis time for all further experiments.

Amino Acid	Expected	24 Hours	36 Hours	48 Hours
Asx	12	12.2	12	11.9
Thr	11	11.4	10.8	10.9
Ser	10	8.9	8.5	8.4
Glx	19	21.9	21.0	20.9
Pro	8	9.4	10.4	9.9
Gly	9	11.7	11.8	11.6
Ala	19	19.9	19.1	19.5
Cys	4	1.5	1.2	1.1
Val	11	10.5	10.0	10.4
Met	1	0.8	0.9	0.7
Ile	5	4.6	4.6	4.7
Leu	23	21.2	18.7	18.8
Tyr	4	4.2	4.8	4.9
Phe	4	4.3	4.5	4.4
His	2	2.8	2.9	2.9
Lys	8	8.7	8.6	8.7
Arg	13	13.5	13.1	13.3

Table 3.1 Amino Acid Analysis of Resin Bound dhEPO

3.5.1 Introduction of the Tbfmoc Group

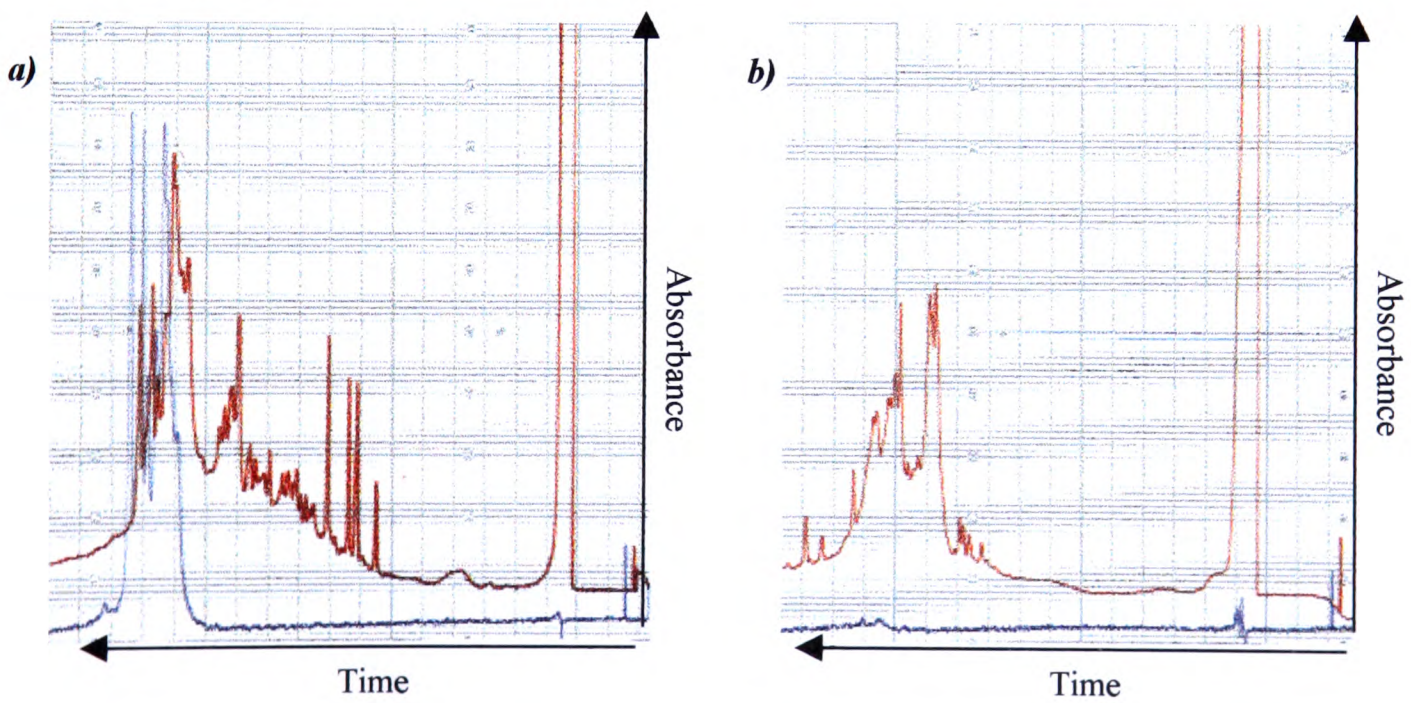
As detailed for Endothelin-1 (**Section 2.9.2**), labelling the free *N*-terminus of a peptide or protein with the Tbfmoc group provides a useful aid for purification. Prior to the introduction of the Tbfmoc moiety, the resin bound Fmoc protected EPO

was subjected to a further capping cycle, to ensure any free amino groups were acylated. Cleavage of the Fmoc group with piperidine enabled introduction of the Tbfmoc moiety as the chloroformate.

Cleavage of the protein from the resin and side chain protecting groups was then achieved in acidiolytic conditions. Scavengers were added to this solution to mop-up carbocations generated from removal of the side chain protecting groups to prevent them reattaching to other reactive functional groups in the protein backbone. In order to minimise the length of time the protein was subjected to the acid conditions, a trial cleavage was carried out on a small sample of resin to determine the optimum cleavage time. Samples were removed at hourly intervals and analysed by HPLC. Cleavage was observed to be complete after 4 hours.

3.5.2 Tbfmoc Affinity Purification on Charcoal

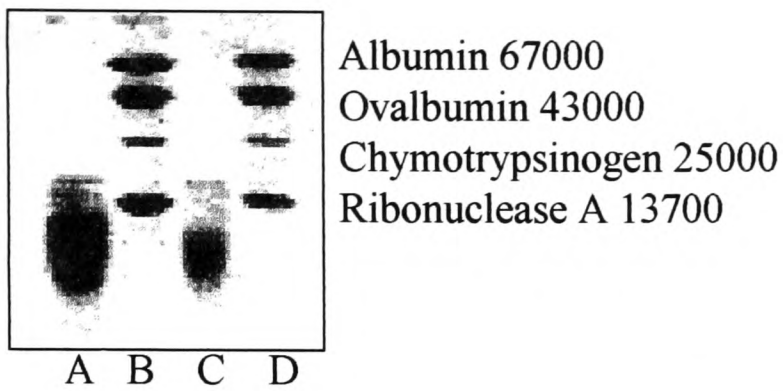
Affinity purification of the Tbfmoc labelled dhEPO was carried out using charcoal as the solid support. After cleavage from the charcoal, the material was passed down a Sephadex G-50 gel filtration column to remove the piperidine salts. Initial analysis by HPLC indicated that the purification had been successful in removing the non-labelled material.



(λ_1 (red) = 214nm, λ_2 (blue) = 364nm)

Figure 3.5 HPLC of dhEPO a) before and b) after Tbfmoc Affinity Purification

Subsequent analysis of the material by SDS-PAGE indicated that, although the correct material appeared to be present, the majority of the material was of a lower molecular weight.



(Lane A, crude Tbfmoc-dhEPO, Lane C dhEPO after affinity purification, Lanes B & D Standards)

Figure 3.6 SDS-PAGE of Crude and Purified dhEPO

This suggested that low molecular weight material had also been tagged with the Tbfmoc group. During the synthesis of the protein, it is possible that the swelling properties of the resin change. As a result, some of the free amino groups become buried and are temporarily unavailable for coupling. Eventually the swelling of the resin again changes and these groups are free to participate in the synthesis. The resulting protein has part of the sequence missing and is referred to as a deletion sequence.

3.5.3 Conclusions

Although the synthesis of dhEPO appeared initially to be successful, closer analysis during purification suggested that a number of deletion sequences had been formed. SDS-PAGE proved useful in assessing the overall purity of the material produced. Since only a very small quantity of the desired material was observed, it was decided not to continue with the purification. During the purification problems were also encountered with the solubility of the crude synthetic material, indeed 8M urea did not completely solubilise the protein.

3.6 Synthesis of dhEPO Employing the Picolyl Group for Cysteine Protection

One problem encountered in the purification of dhEPO was the low solubility of the protein in a number of solvents. It is known that synthetic proteins containing cysteine residues can aggregate due to interactions between the free thiol groups.[31] Although there is no evidence that this problem had occurred, it was reasonable to assume that purification of the protein would be aided if the thiol group remained protected at this stage. As a result, synthesis of the protein was repeated using the acid and base stable picolyl protecting group for all four cysteine residues. This was also expected to aid the solubility of the protein.

Recent research has suggested that the initial loading of the resin is crucial for the success of a large synthesis.[32] It has been noted that the synthesis of dhEPO proceeds with much greater efficiency when the initial functionality of the resin is no greater than 0.1mmol/g.

The synthesis of dhEPO(Pic) was carried out on a 0.1mmol scale using Fmoc-Arg(Pbf) functionalised Wang resin, loaded with an initial functionality of 0.1mmol/g. From the deprotection profile shown below it can be seen that the synthesis appeared to proceed more efficiently, indeed the percentage coupling never dropped below 95% (**Figure 3.7**).

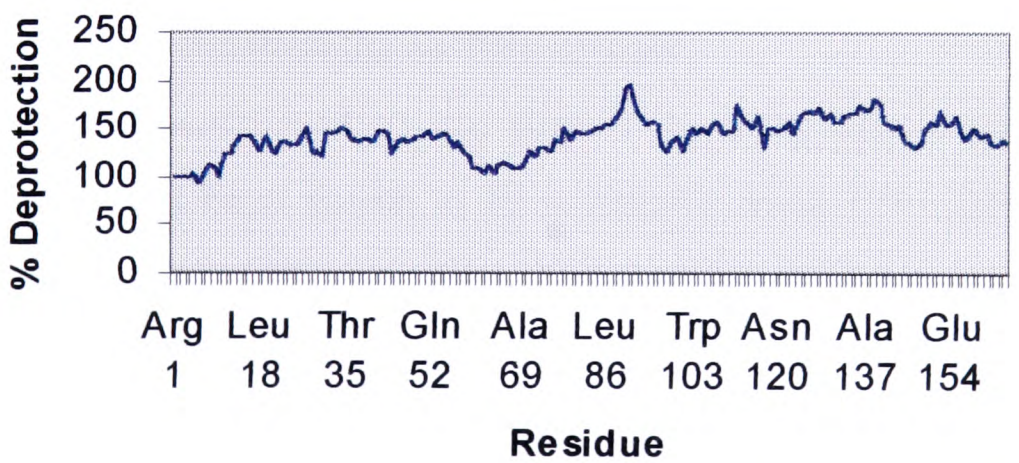


Figure 3.7 Fmoc Deprotection Profile of dhEPO(Pic)

3.6.1 Introduction of the Tbfmoc Moiety

Although used with great success, introduction of the Tbfmoc moiety *via* the highly reactive chloroformate can present some problems. It has been proposed that the Tbfmoc moiety may associate with some of the side chain functional groups of the peptide or protein. As a result, coupling solely to the free amino group of a peptide or protein cannot be guaranteed. Introduction of the Tbfmoc protected derivative of the final amino acid should overcome this potential problem. Extensive studies have been carried out to develop an efficient route to the synthesis of these N^{α} protected amino acids. Tbfmoc-Ala-OH (**3.1**) was synthesised according to the procedure of Lorthioir.[33]

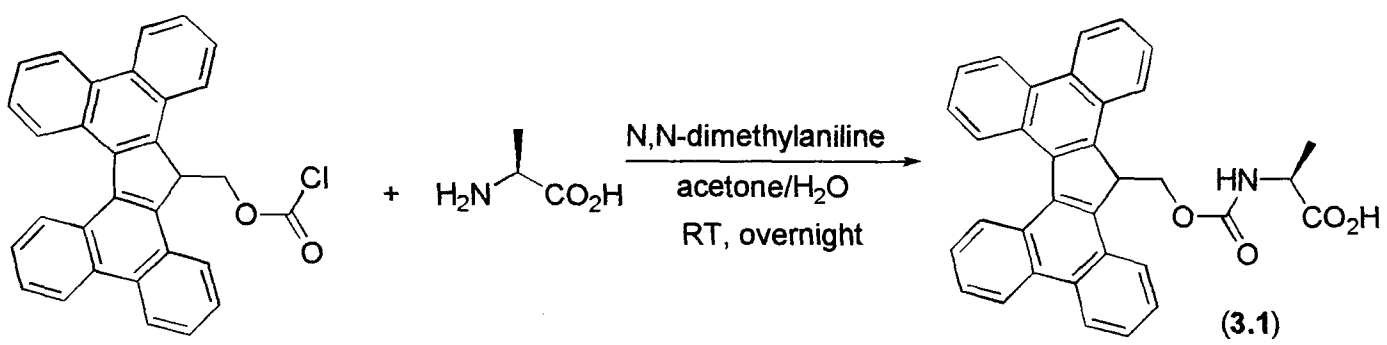


Figure 3.8 Synthesis of Tbfmoc-Ala-OH

Coupling of the Tbfmoc protected amino acid is facilitated *via* the active ester. Subsequent cleavage of Tbfmoc-dhEPO(Pic) from the resin and side chain protecting groups was achieved under acidiolytic conditions as detailed earlier.

3.6.2 Tbfmoc Affinity Purification on Polystyrene

Purification of Tbfmoc-dhEPO(Pic) was carried out using chromatographic grade polystyrene, which has been shown to be a superior support to charcoal.[34] The uniform structure of the polystyrene ensures that there are no fine particles and thus a pre-washing step can be avoided. The rigid nature of the polystyrene also facilitates washing by a simple filtration step.

Once cleaved from the polystyrene-Tbfmoc group, dhEPO(Pic) was desalted by gel filtration. Although some purification of the material appeared to have been

successful, the yield of material recovered after the gel filtration step was less than 5%. As a result it was felt that this purification step was not applicable for this particular protein. It is possible that non-specific absorptions to the polystyrene are responsible for this disappointing yield. Such absorptions could occur *via* the side chains of the aromatic amino acids or through hydrophobic interactions with the protein and the polystyrene. Since the protein is unlikely to have any defined secondary or tertiary structure at this time, it is possible that the hydrophobic amino acids are still exposed to solvent rather than buried within the core of the protein.

3.6.3 Sephadex G75 Gel Filtration of dhEPO(Pic)

Size exclusion gel filtration is a powerful tool for the purification of proteins, molecules being separated according to size. The gel is composed of a porous matrix, which is permeable to small molecules. These become trapped, hindering their progress down the column. As the molecules increase in size they are trapped to a lesser degree and so pass through the column with little difficulty.

A number of factors affect the efficiency of gel filtration. The length of the column is significant since it affects both the retention time of the compound and also the degree of resolution. Generally, the volume of the sample load should be less than 1% of the column volume to ensure maximum resolution. A variety of solvents can be used in gel filtration, the main requirements being that they are compatible with the chemical structure of the gel medium and that the protein be readily soluble in the eluent.[35]

Sephadex G-75 Superfine grade was chosen for purification of dhEPO(Pic). Sephadex is formed from the cross-linking of dextran and epichlorohydrin and displays excellent swelling properties in aqueous media. Sephadex G-75 has a fractionation range of 3000-70000 and so is ideal for purification of dhEPO(Pic) which has a calculated mass of 18760.14. Greater resolution is obtained with the Superfine grade of gel since a slow flow rate is maintained. Previous studies by

Patterson have shown Sephadex G-75 Superfine grade to be an ideal medium for purification of chemically synthesised human interferon-γ.[36]

DhEPO(Pic) was applied to the column in 6M Urea containing 0.2M NaCl and 0.1M phosphate buffer pH 7.5. The protein was then eluted using 20% v/v AcOH/H_2O. Fractions were collected at hourly intervals and assessed for protein content by UV. A sample of all fractions containing protein was removed and lyophilised. SDS-PAGE of this material enabled detection of the material containing the protein of interest. Generally the first four or five fractions containing protein were found to contain dhEPO(Pic) by SDS-PAGE. The efficiency of the column can be seen in the SDS-PAGE of the fractions collected.

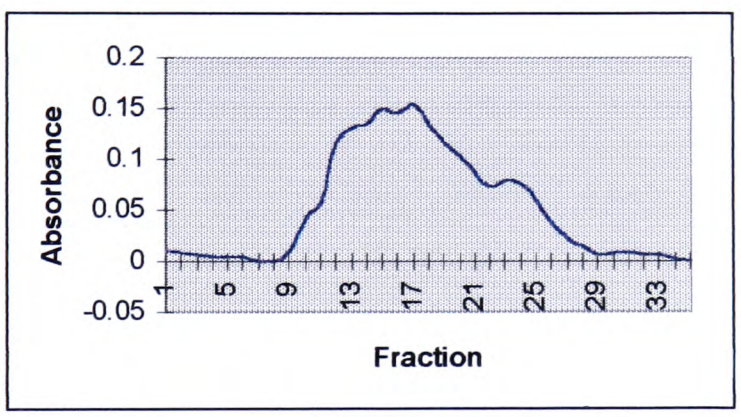


Figure 3.9 UV Analysis of dhEPO(Pic) After 1st Pass down Gel Column

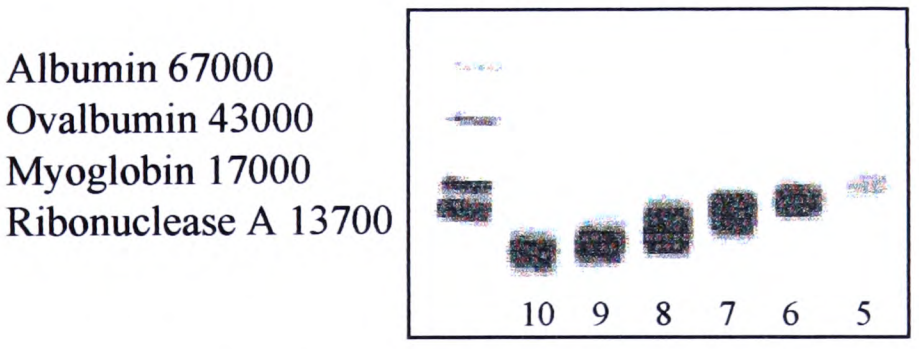


Figure 3.10 SDS-PAGE Analysis of dhEPO(Pic) After 1st Pass down Gel Column

(Each lane in gel corresponds to a different fraction eluted from the gel column)

All dhEPO(Pic) containing fractions were pooled, lyophilised and subjected to a further round of size-exclusion gel filtration. The detection and analysis procedure detailed above was repeated. From the UV trace of the fractions the emergence of the purified protein is clearly visible in fractions 9-11 (**Figure 3.11**).

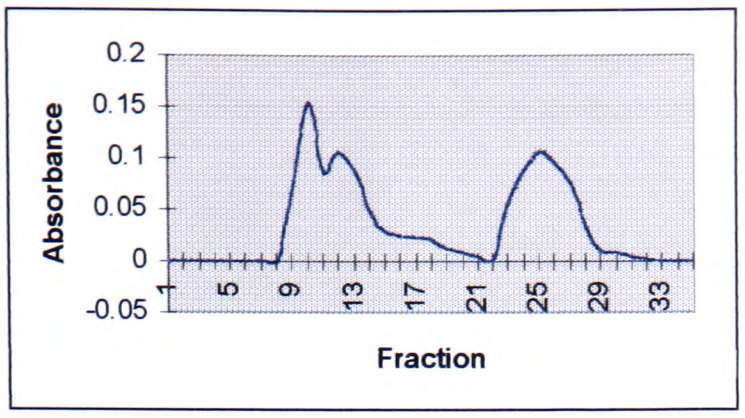


Figure 3.11 UV Analysis of dhEPO(Pic) After 2nd Pass down Gel Column

(Peak in fractions 23 – 28 due to urea eluting from column)

This material was again pooled and lyophilised. A final application to the gel filtration column generated the pure dhEPO(Pic).

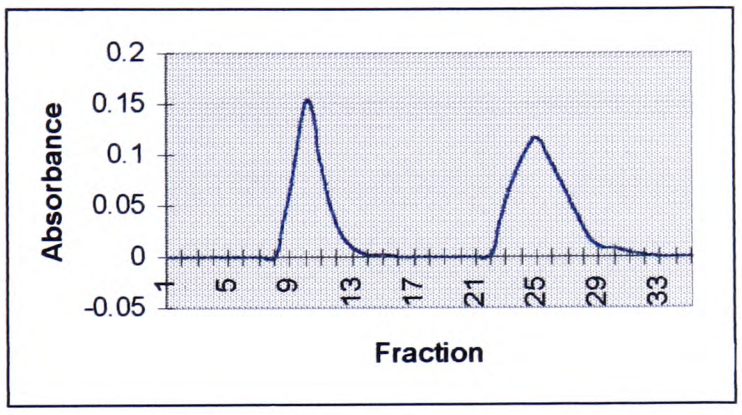


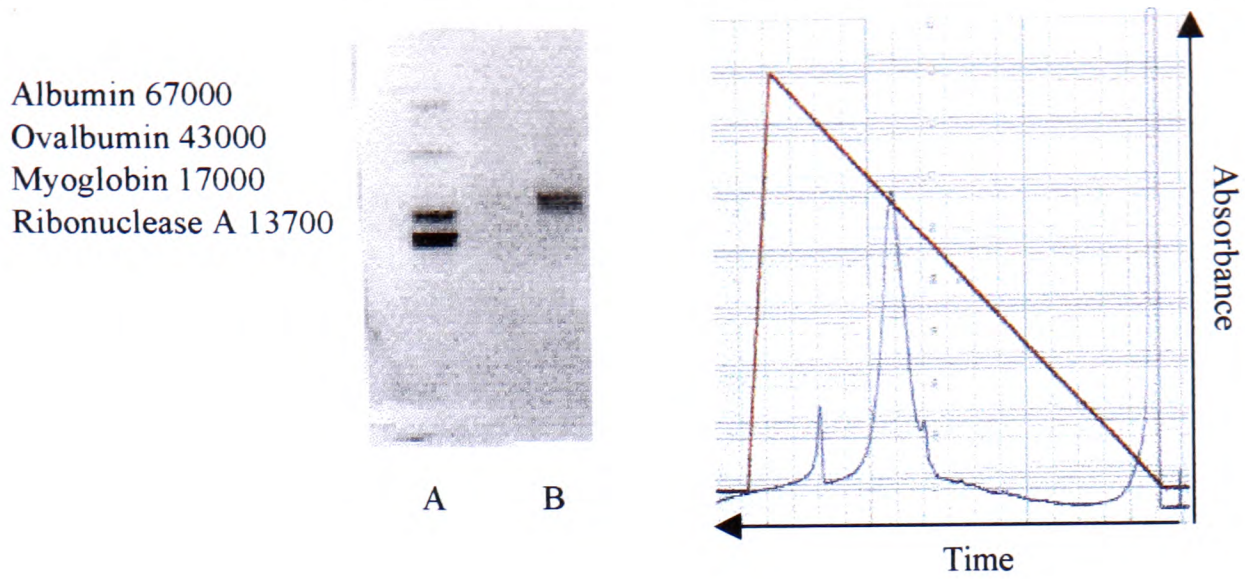
Figure 3.12 UV Analysis of Purified dhEPO(Pic)

3.6.4 Characterisation of dhEPO(Pic)

3.6.4.1 SDS-PAGE and HPLC

Analysis of the purified protein by SDS-PAGE confirmed the presence of the single band material. In common with other synthetic proteins, the band appears to be slightly broader than recombinant proteins, although this is possibly due to traces of salt still present in the sample.

The HPLC trace confirms the presence of a single component. The broad peak is again characteristic of synthetic proteins and possibly accounts for the denatured state of the protein.



(Lane A Standards, Lane B Purified dhEPO(Pic)

Figure 3.13 SDS-PAGE and HPLC Trace of dhEPO(Pic)

3.6.4.2 Tryptic Digest

Since MALDI-ToF analysis on the purified protein was unsuccessful, the protein was subjected to enzymatic cleavage and the fragments produced analysed by MALDI-ToF mass spectrometry. Trypsin was chosen as the proteolytic enzyme, cleavage occurring at the *C*-termini of lysine and arginine residues. The cleavage was complete after two hours. All possible cleavage sites appear to have been cut by the enzyme, only 12% of the sequence remaining unidentified by mass spectroscopy. The results for the cleavage are detailed in **Table 3.2**.

Fragment	**Molecular Weight (MH^+)**	**Observed Mass**	**Fragment**	**Molecular Weight (MH^+)**	**Observed Mass**
Leu^{5}-Arg^{10}	797.36	797.49	Ser^{104}-Arg^{110}	803.49	803.16
Tyr^{15}-Lys^{20}	736.42	738.96	Ala^{111}-Lys^{116}	587.35	588.99
Glu^{21}-Lys^{45}	2871.2	2868.6	Glu^{117}-Arg^{131}	1465.76	1465.21
Val^{46}-Lys^{52}	927.47	969.61 (K^+)	Thr^{132}-Arg^{139}	924.48	923.07
Met^{54}-Arg^{76}	2526.34	2548.76 (Na^+)	Val^{144}-Arg^{150}	898.45	922.99 (Na^+)
Gly^{77}-Lys^{97}	2359.2	2359.52	Leu^{155}-Arg^{162}	1003.42	1005.16
Ala^{98}-Arg^{103}	602.36	602.57			

Table 3.2 Mass Spectra Data from Tryptic Digest

3.6.4.3 *N*-Terminal Sequencing[37]

Tryptic digest was successful in identifying the majority of the polypeptide chain, however there was no evidence for the four *N*-terminal residues. In order to confirm that the synthesis had proceeded to this point a sample of the protein was analysed by *N*-terminal sequencing. Edman degradation is a powerful technique, which enables sequential identification of the amino acids from the *N*-terminus. Edman's reagent (phenyl isothiocyanate) reacts with the *N*-terminal amino acid to produce a derivative, which can be released for identification without hydrolysis of the rest of the polypeptide chain. Five rounds of Edman degradation were performed satisfactorily on dhEPO(Pic), confirming the presence of the *N*-terminal sequence of the protein.

3.6.4.4 Isoelectric Focusing

The isoelectric point (pI) of a protein is defined as the pH where the individual charges of the amino acids cancel out and the net charge is zero. Isoelectric focusing (IEF) is a non-denaturing technique, which enables the pI of a protein to be determined. By subjecting the protein to an electric current within a pH gradient, it will migrate to the point where the net charge is zero. The pH gradient can be established by the use of ampholytes. These are small, charged buffer molecules which quickly migrate within the cell, establishing a pH gradient increasing from anode to cathode.

The protein was loaded into the focusing chamber of the Bio-Rad Rotofor® cell along with ampholytes (pH range 3.0 – 10.0). The system was subjected to a constant power and the increase in voltage noted over time. Once the voltage stabilised, the protein was judged focused at its isoelectric point.

After focusing the individual fractions were harvested and the pH of each measured. Each fraction was analysed for protein content by HPLC.

Fraction	pH	Fraction	pH	Fraction	pH	Fraction	pH
1	3.5	**6**	5.5	**11**	7.5	**16**	8.5
2	4.0	**7**	6.0	**12**	7.5	**17**	9.0
3	4.5	**8**	7.0	**13**	8.0	**18**	9.5
4	4.5	**9**	X	**14**	X	**19**	10.0
5	5.0	**10**	X	**15**	8.0	**20**	10.0

(X denotes fractions not collected)

Table 3.3 IEF of dhEPO(Pic)

The protein was found in fractions 19 and 20 by HPLC analysis, indicating a pI of 10.0. The pI of dhEPO has been measured as 8.75,[38] therefore it is reasonable to assume that introduction of the basic picolyl protecting group on all four cysteine residues will cause an increase in the pI of the protein.

3.6.4.5 Amino Acid Analysis

A sample of the protein was hydrolysed in 6N HCl for 24 hours and the amino acid composition analysed. The results obtained are in reasonable agreement with the predicted values, confirming that the purification protocol has been successful.

Amino Acid	Expected	24 Hours
Asx	12	11.8
Thr	11	9.7
Ser	10	8.6
Glx	19	19.8
Pro	8	8.2
Gly	9	9.2
Ala	19	18.9
Val	11	11.4
Met	1	1.2
Ile	5	5.2
Leu	23	22.9
Tyr	4	3.6
Phe	4	4.0
His	2	2.2
Lys	8	7.9
Arg	13	11.9

Table 3.4 Amino Acid Analysis of dhEPO(Pic) after Purification

3.6.4.6 Molecular Weight Determination by FPLC

The molecular weight of the synthetic dhEPO(Pic) was analysed by size exclusion gel filtration using a Superdex™ column.[39] The column was calibrated with a series of molecular weight standards (Ribonuclease A, MWt 13700; Chymotrypsinogen A, MWt 25000; Ovalbumin, MWt 43000) and the void volume calculated using Dextran Blue. Thus, the elution volume parameter K_{AV} can be calculated from the equation;

$$K_{AV} = (Ve - Vo) / (Vt - Vo)$$

where Ve is the elution volume, Vo the column void volume and Vt the column bed volume. Plotting K_{AV} against log MWt enables calculation of the molecular weight of the synthetic material from the least squares line.

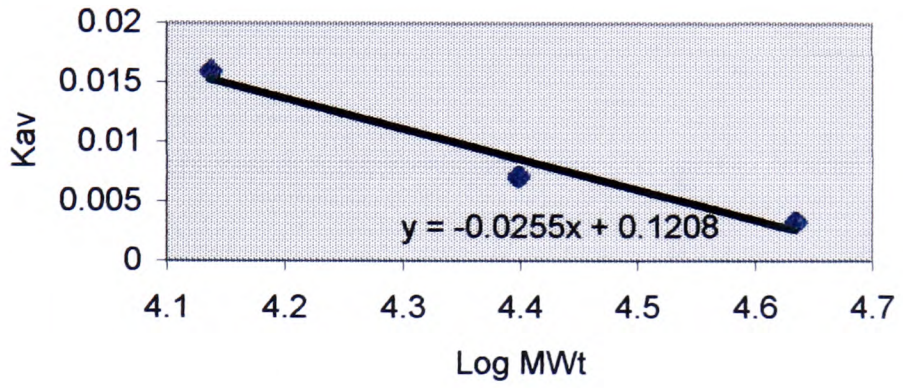


Figure 3.14 Calibration of the Gel Filtration Column

The calculated molecular weight of dhEPO(Pic) is 18.81 kDa. This is in reasonable agreement with the expected value of 18.76 kDa.

3.5 Cleavage of the Picolyl Protecting Groups

To be considered a useful method for synthesis and purification of dhEPO, cleavage of the picolyl protecting groups must be rapid and quantitative. Previous work with oxytocin and endothelin-1 enabled picolyl group cleavage to be monitored by HPLC. Unfortunately, the broad HPLC peak of dhEPO(Pic) means that analysis of the picolyl group cleavage is unlikely to be possible by this method. Previous work with the endothelin system, which also contains four picolyl groups, has shown that quantitative cleavage can be accomplished in 1 hour.

The protein was subjected to the picolyl group cleavage conditions for 2 hours. The time was extended slightly to ensure complete deprotection. Semi-preparative HPLC was employed to separate dhEPO from the zinc salts formed during the reaction.

The material obtained from the reaction gave a positive test with Ellmans' Reagent, indicating that the material did indeed contain free thiol groups. In order to demonstrate that the cleavage had gone to completion a sample of the material was subjected to amino acid analysis. As noted in **Chapter 2**, the picolyl protected cysteine is stable to acid hydrolysis, eluting near histidine in the amino acid analysis chromatogram. Cleavage of the picolyl groups was confirmed from the amino acid analysis chromatogram.

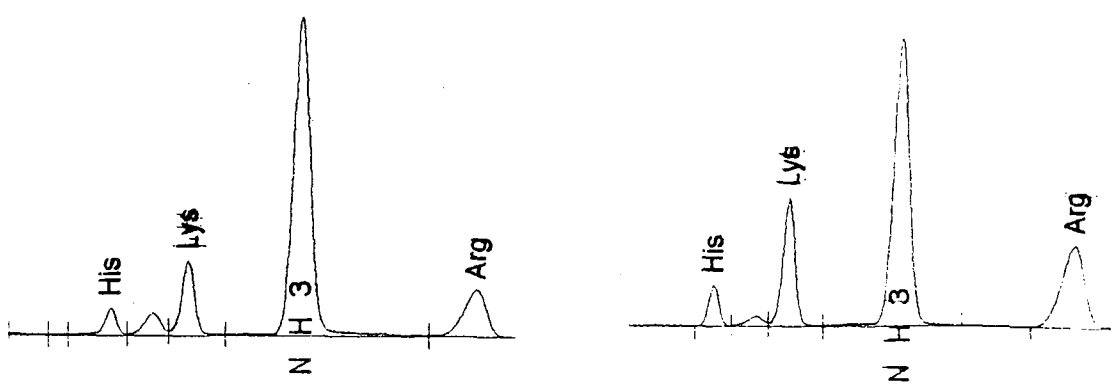


Figure 3.15 Amino Acid Analysis Chromatogram before and after picolyl group cleavage

3.7 Conclusions and Recommendations

The synthesis and purification of dhEPO(Pic) has been successful. Protection of the cysteine residues with the picolyl group aided the solubility and hence the purification of the protein. Consistent with previous results, the synthesis of the protein was found to proceed with much greater efficiency when an initial resin loading of no greater than 0.1mmol/g was employed.

A short efficient purification protocol has been described, which should be readily amenable to scale up. Cleavage of the picolyl protecting groups presented no problems, complete cleavage occurring within 2 hours.

To carry out any biological assays, the protein must be folded into the biologically active form. In nature it is believed that proteins are maintained in a partially folded conformation until they reach their final cellular location. Here folding is achieved in the presence of specific chaperones or folding catalysts.[40] Molecular chaperones such as GroES and GroEL act on the unfolded protein causing release of either the fully folded material or a structural intermediate that completes folding after release from the cell.[41,42] Proteins containing disulfides require catalytic oxidation by a member of the thioredoxin family of enzymes. These include glutaredoxin, PDI, DsbA and DsbC.[43,31] Production of these enzymes by recombinant methods should enable studies into the folding of the synthetic material. Recombinant dhEPO, which was expressed in *E. coli,* has been successfully folded using copper sulfate mediated oxidation.[12] With a combination of the above methods it is likely that folding of the synthetic material could be achieved, leading to structural and biological activity studies.

1. P. Carnot and C. Deflandre, *C. R. Hebd. Seanecs. Acad. Sci. Ser. M.*, 1906, **3**, 384.

2. P. Manavalan, D. L. Swope and R. M. Withy, *J. Prot. Chem.*, 1992, **11**, 321.

3. A. J. Erslev, *Blood*, 1955, **10**, 954.

4. T. Miyake, C. K.-H. Kung and E. Goldwasser, *J. Biol. Chem.*, 1977, **252**, 5558.

5. K. Jacobs C. Shoemaker, R. Rudersdorf, S. D. Neill, R. J. Kaufman, A. Mufson, J. Seehra, S. S. Jones, R. Herrick, E. F. Fritsch, M. Kawakita, T. Shimizu and T. Miyake, *Nature*, 1985, **313**, 806.

6. F.-K. Lin, S. Suggs, C.-H. Lin, J. K. Browne, R. Smalling, J. C. Egrie, K. K. Chen, G. M. Fox, F. Martin, Z. Stabinsky, S. M. Badrawi, P.-H. Lai and E. Goldwasser, *Proc. Natl. Acad. Sci. USA*, 1985, **22**, 7580.

7. J. C. Egrie, T. W. Strickland, J. Lane, K. Aoki, A. M. Cohen, R. Smalling, G. Trail, F. K. Lin, J. K. Browne and D. K. Hines, *Immunobiology*, 1986, **172**, 213.

8. F.-K. Lin United States Patent 4,703,008.

9. R. T. Means, *Hematology/Oncolocy Clinics of North America*, 1994, **8**, 933.

10. D. H. Henry, *Hematology/Oncolocy Clinics of North America*, 1994, **8**, 961.

11. M. A. Rency, H. A. Scoble and Y. Kim, *J. Biol. Chem.*, 1987, **262**, 17156.

12. J.-P. Boissel, W-R. Lee, S. R. Presnell, F. E. Cohen and H. F. Bunn, *J. Biol. Chem.*, 1993, **268**, 15983.

13. T. R. J. Lappin, P. C. Winter, G. E. Elder. C. M. M^{c}Hale, V. H. Hodges and J. M. Bridges, *Annals. New York Acad. Sci.*, 1994, **718**, 191.

14. S. Dubé, J. W. Fisher and J. S. Powell, *J. Biol. Chem.*, 1988, **263**, 17516.

15. K. B. Linsley, S.-Y. Chan, S. Chan, B. B. Reinholt, P. J. Lisi and V. N. Reinholt, *Analytical Biochemistry*, 1994, **219**, 207.

16. M. S. Dordal, F. F. Wang & E. Goldwasser, *Endocrinology*, 1985, **116**, 2293.

17. M. N. Fukuda, H. Sasaki, L. Lopez and M. Fukuda, *Blood*, 1989, **73**, 84.

18. E. Tsuda, G. Kawanishi, M. Ueda, S. Masuda and R. Saski, *Eur. J. Biochem.*, 1990, **188**, 405.

19. J. L. Spivak, *Nephron*, 1989, **52**, 289.

20. D. Wen, J.-P. Boissel, M. Showers, B. C. Ruch and H. F. Bunn, *J. Biol. Chem.*, 1994, **269**, 22839.

21. D. Wen, J.-P. Boissel, T. E. Tracy, R. H. Gruninger, L. S. Mulcahy, J. Czelusniak, M. Goodman and H. F. Bunn, *Blood*, 1993, **82**, 1507.

22. J. C. Cheetham, D. M. Smith, K. H. Aoki, J. L. Stevenson, T. J. Hoeffel, R. S. Syed, J. Egrie and T. S. Harvey, *Nature Structural Biology*, 1998, **5**, 861.

23. R. S. Syed, S. W. Reid, C. Li, J. C. Cheetham, K. H. Aoki, B. Lui, H. Zhan, T. D. Osslund, A. J. Chirino, J. Zhang, J. Finer-Moore, S. Elliott, K. Sitney, B. A. Katz, D. J. Matthews, J. J. Wendoloski, J. Egrie and R. M. Stroud, *Nature*, 1998, **395**, 511.

24. S. S. Jones, A. D. D'Andrea, L. L. Haines and G. C. Wong, *Blood*, 1990, **76**. 31.

25. H. Youssoufian, G. Longmore, D. Neumann, A. Yoshimura and H. F. Lodish, *Blood*, 1993, **81**, 2223.

26. N. Robertson and R. Ramage, *J. Chem. Soc., Perkin Trans. I.*, 1999, 1015.

27. L. C. Draffan, *PhD Thesis*, University of Edinburgh, 1996.

28. G. H. Morton, *PhD Thesis,* University of Edinburgh, 1997.

29. N. Robertson, *PhD Thesis,* University of Edinburgh, 1996.

30. K. M. Otteson, R. L. Nobel, P. D. Hoeprich Jr., K. T. Shaw and R. Ramage, *Applied Biosystems Res. News*, June 1993.

31. A. Zapun, D. Missiakas, S. Raina and T. E. Creighton, *Biochemistry*, 1995, **34**, 5075.

32. E. Saurez, *Personal Communication*, University of Edinburgh.

33. O. Lorthioir, *Post-Doctoral Research Report*, University of Edinburgh, 1997.

34. C. Jamieson, and J. A. Patterson, *Unpublished Results*, University of Edinburgh.

35. In 'Gel Filtration, Principles and Methods', *Pharmacia Biotec*. 1993.

36. J. Patterson, *PhD Thesis*, University of Edinburgh, 1999.

37. A. D. Cronshaw, Welmet Sequencing, University of Edinburgh.

38. N. Imai, A. Kawamura, M. Higuchi, M. Oh-eda, T. Orita, T. Kawaguchi and N. Ochi, *J. Biochemistry*, 1990, **107**, 532.

39. In 'Protein Purification. Principles, high resolution methods and applications', pp63 – 106. (ed. J.-C. Janson, L. Rydén), VCH Publishers Inc., New York, Weinhem, Cambridge, 1989.

40. D. Missiakas, C. Georgopoulos and S. Raina, *The EMBO Journal*, 1994, **13**, 2013.

41. M. Mayhew, A. C. da Silva, J. Martin, H. Erdjument-Bromage, P. Tempst and F. –U. Hartl, *Nature*, 1996, **379**, 420.

42. J. S. Weissman, H. S. Rye, W. A. Fenton, J. M. Beechem and A. L. Horwich, Cell, 1996, **84**, 481.

43. J. L. Martin, J. C. A. Bardewell and J. Kuriyan, *Nature*, 1993, **365**, 464.

CHAPTER 4

Studies Towards the Convergent Synthesis of dhErythropoietin

4.1 Introduction

Chapter 3 detailed the successful stepwise synthesis and purification of dhEPO. However, purification of large proteins obtained by chemical synthesis is hampered by the presence of deletion sequences. Deletions may be formed which lack only one or two residues. These sequences are likely to be almost impossible to identify and even harder to separate from the correct material. One way to overcome this problem is to adopt a convergent approach to the synthesis. Thus, a number of smaller peptide fragments are synthesised and coupled to generate the desired protein. The advantage of this approach being that the synthesis and purification of the smaller peptide fragments is simpler than that of the complete protein. Moreover, the differences between the ligated material and the individual fragments will be more pronounced easing identification and separation of impurities.

As the goals of chemical protein synthesis become ever more ambitious, new methodology must be developed. Fragment coupling is perhaps the most promising method for accessing proteins of greater than 200 residues. One notable technique is that of native chemical ligation developed by Kent and co-workers.[1,2,3] This has successfully been applied in the synthesis of HIV-1 protease[4] and Turkey Ovomucoid Third domain.[5] Nevertheless, the traditional method of fragment coupling, proceeding *via* the azide intermediate, seems to offer most scope.

4.2 The Azide Method[6,7]

Curtius introduced the azide method in 1902 for peptide bond formation.[8] The methyl or ethyl ester of a peptide could be converted into the peptide hydrazide on treatment with hydrazine hydrate. Subsequent reaction of this material with an

organic nitrite produced a reactive azide intermediate, which could undergo coupling with the free amino group of another amino acid. Due to the instability of azides they are generally not isolated but used immediately *in situ*. Conversion of peptide hydrazides to peptide azides was studied exhaustively by Honzel and Rudinger.[9] They found the reaction to proceed most efficiently at low temperatures in an aprotic organic solvent such as DMF. This remains the most commonly used method for conversion of peptide hydrazides to azides.

For many years it was thought that the azide method was completely free of racemisation. However, numerous studies have found some degree of racemisation results when the peptide azide is exposed to strong bases.[10,11,12,13] Despite this the azide method retains popularity since it presents minimal risk of racemisation in optimum conditions. The high reactivity of the azide intermediate requires that reactions be carried out at low temperatures. Thus, problems with the solubility of the peptide segments are introduced.

4.3 Transfer Active Ester Condensation

Extensive studies have been carried out to find an efficient alternative to the azide method. Transfer Active Ester Condensation (TAEC) has been developed from modification of the azide method.[14,15] This method retains all the advantages of the azide method whilst eliminating the need for low temperatures and hence the associated peptide solubility problems. As in the traditional azide method the peptide hydrazide (**4.1**) is treated with an organic nitrite to form the peptide azide (**4.2**). However, when the reaction is carried out in the presence of a large excess of HOCt the peptide azide is trapped *in situ* to produce the peptide-OCt derivative (**4.3**), **Figure 4.1**. This derivative can then react with the amino terminus of a second peptide fragment to produce the coupled material with a peptide bond at the site of union (**4.4**).

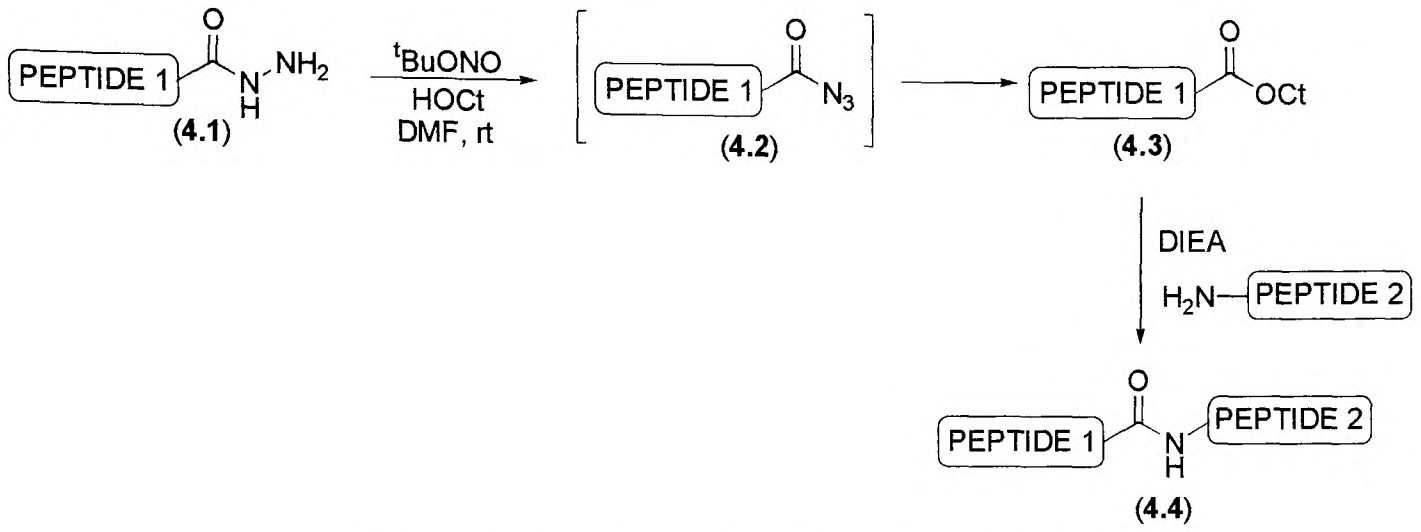


Figure 4.1 Transfer Active Ester Condensation

Peptide-OCt derivatives are more stable than the corresponding peptide azides with greater solubility also observed. Furthermore, due to the acidity of HOCt, addition of a mineral acid is not required during the formation of the azide intermediate.

4.4 Outline of Research

It was proposed that the synthesis of dhEPO be repeated using TAEC. The preferred coupling sites are at small non-functionalised residues such as glycine, alanine, leucine and valine. Analysis of the dhEPO sequence identified four possible ligation sites, which split the protein into five fragments of reasonable size, ranging from 26 to 39 residues.

Ala Pro Pro Arg Leu Ile Cys Asp Ser Arg Val Leu Glu Arg Tyr Leu
Leu Glu Ala Lys Glu Ala Glu Asn Ile Thr Thr Gly Cys Ala Glu His
Cys Ser Leu Asn Glu Asn Ile Thr Val Pro Asp Thr Lys Val Asn Phe
Tyr Ala Trp Lys Arg Met Glu Val Gly Gln Gln Ala Val Glu Val Trp
Gln Gly Leu Ala Leu Leu Ser Glu Ala Val Leu Arg Gly Gln Ala Leu
Leu Val Asn Ser Ser Gln Pro Trp Glu Pro Leu Gln Leu His Val Asp
Lys Ala Val Ser Gly Leu Arg Ser Leu Thr Thr Leu Leu Arg Ala Leu
Gly Ala Gln Lys Glu Ala Ile Ser Pro Pro Asp Ala Ala Ser Ala Ala
Pro Leu Arg Thr Ile Thr Ala Asp Thr Phe Arg Lys Leu Phe Arg Val
Tyr Ser Asn Phe Leu Arg Gly Lys Leu Lys Leu Tyr Thr Gly Glu Ala
Cys Arg Thr Gly Asp Arg

Figure 4.2 Sequence of dhEPO Potential Ligation Sites Highlighted in Red

4.5 Hydrazide Versus Semicarbazide Linker

Initial research involved modification of the peptide hydrazide. However, there are concerns that, should the azide forming reaction be slow, *bis*-acyl hydrazine (**4.5**) formation may compete with the desired transformation (**Figure 4.3**).

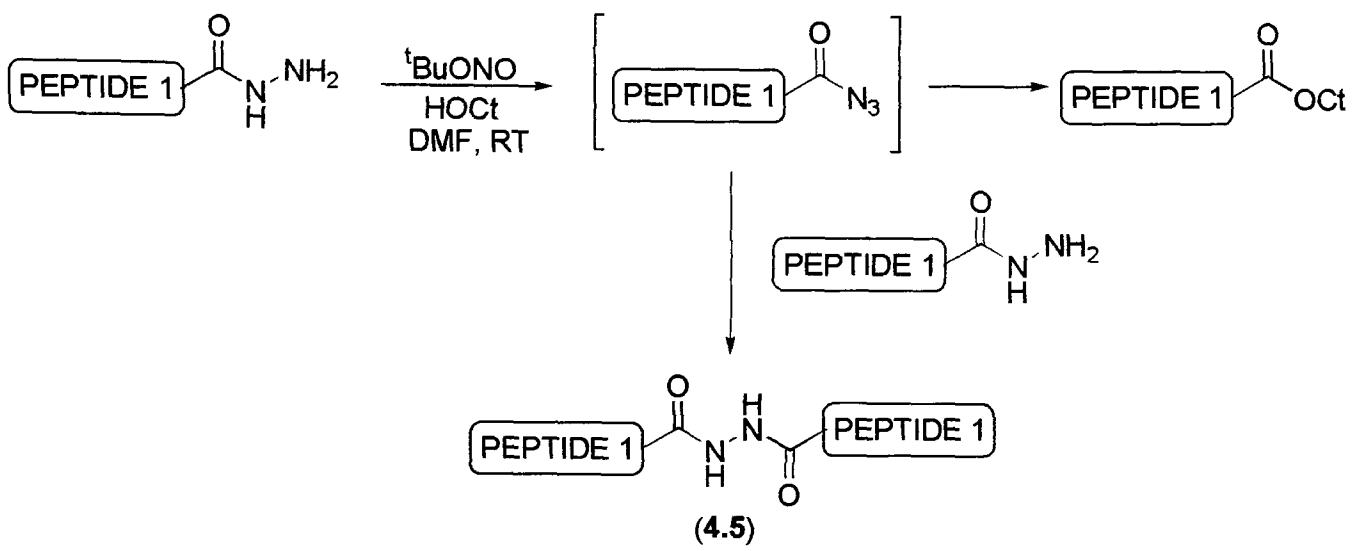


Figure 4.3 Formation of Bis acyl hydrazine

Employing the less reactive peptide semicarbazide (**4.6**) as starting material should prevent this reaction occuring. Recent research suggests that peptide semicarbazides can be efficiently transformed to peptide azides, only requiring a slightly longer reaction time than the corresponding peptide hydrazide.[16]

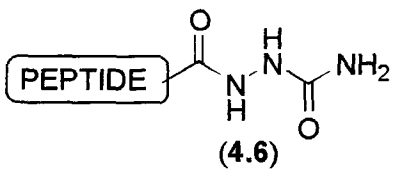


Figure 4.4 Peptide C-Terminal Semicarbazide

Tricyclic amide linker resin (**4.7**) can be modified to generate the semicarbazide resin (**4.9**).[17,18] Peptides are then synthesised on the resin using standard SPPS protocols, subsequent acidiolytic cleavage generating the peptide semicarbazide.

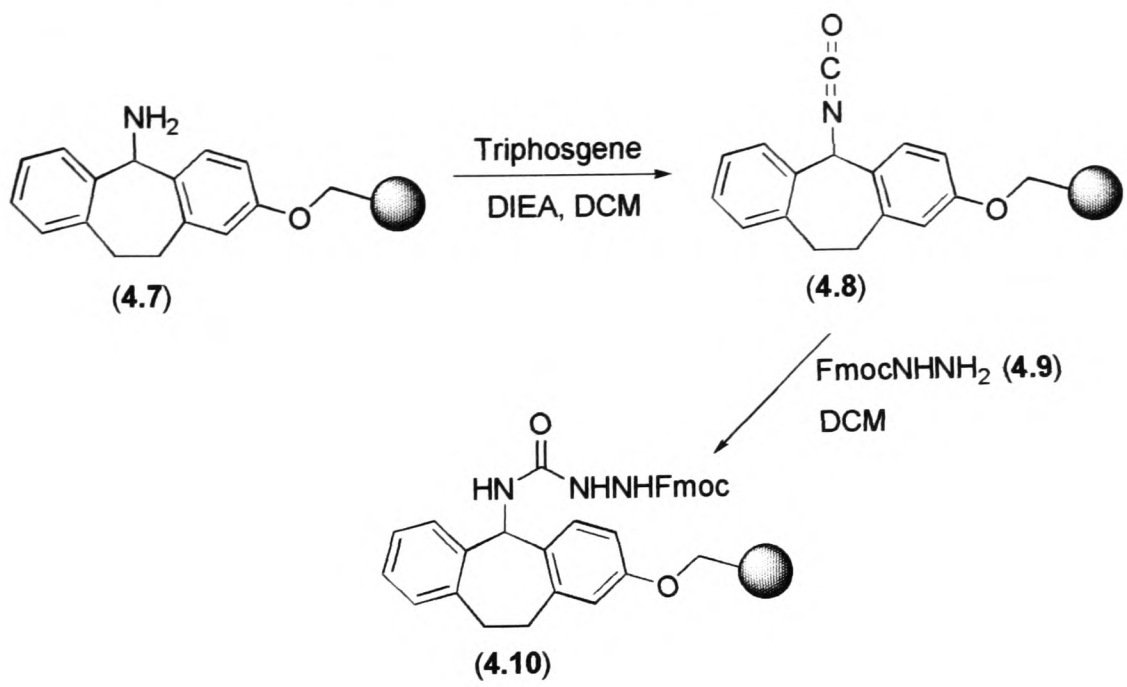


Figure 4.5 Synthesis of Semicarbazide Linker

4.6 Protecting Group Strategies

Perhaps the greatest advantage of TAEC is that the majority of side chain functional groups do not require protection. Consequently unprotected peptide fragments can be ligated, which are inherently more soluble than their protected counterparts. The only side chain functional groups which require protection are the thiol functionality of cysteine residues and the ε-amino group of lysine residues.

The properties of the picolyl protecting group, outlined in **Chapter 2**, make it ideal for cysteine protection. The solublising effect, which the group confers on the peptide chain, should also prove advantageous.

Lysine protection is achieved with the 4-methyloxycarbonyl-4-nitro-[2,6-dioxaspiro-5,5-undecane] (pTnm) group (**4.12**).[19] On treatment with acid the acetal of the pTnm group cleaves to give the 2,2-bishydroxymethyl-2-nitromethyloxycarbonyl (Tnm) group. Cleavage of the Tnm group is readily achieved upon treatment with aqueous

buffer at pH 8.5. Synthesis of the protected lysine derivative can be achieved as outlined in **Figure 4.6**.

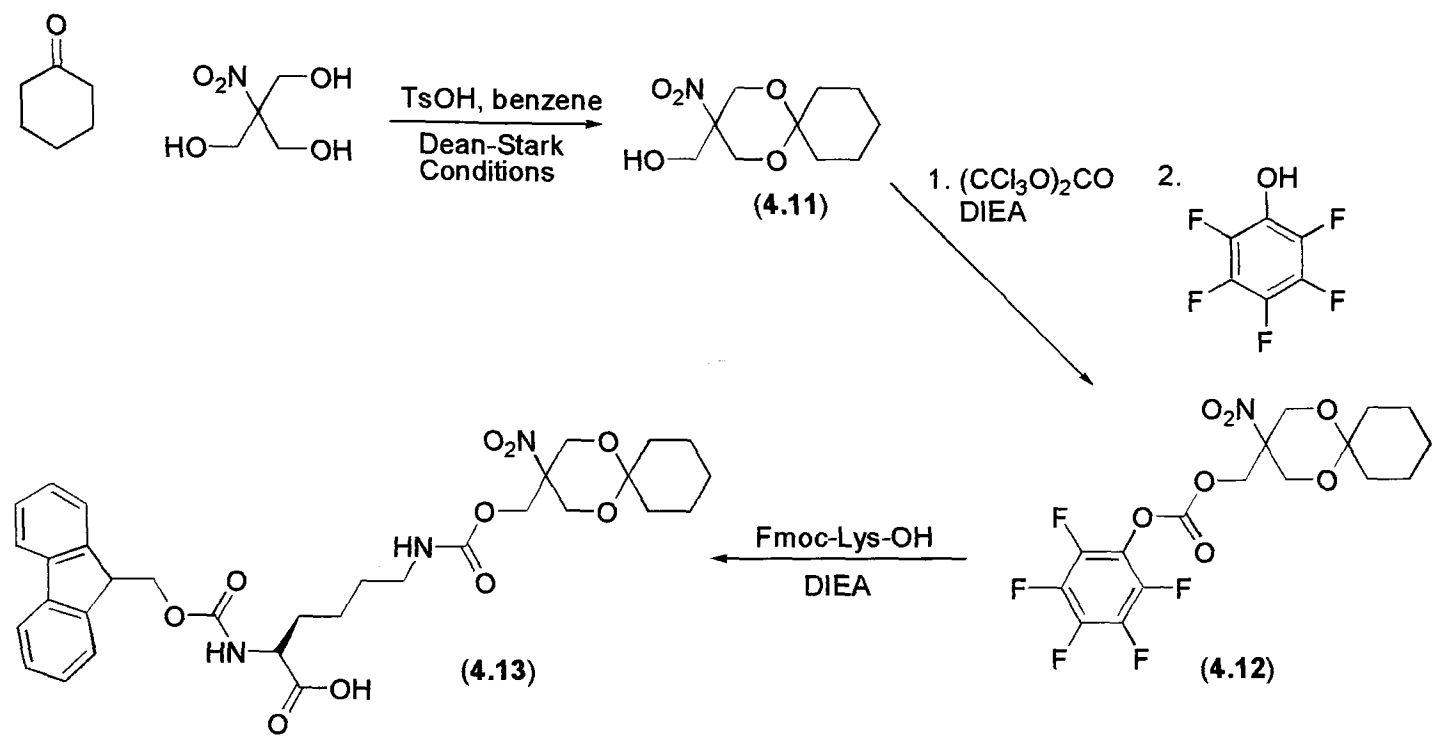


Figure 4.6 Synthesis of Fmoc-Lys(pTnm)-OH

4.7 Synthesis and Purification of the Peptide Fragments

Fragment	MWt (Expected)	MWt (Found)	Amino Acid Analysis (24 hours hydrolysis)
Fragment 5 Ala^{1}-Gly^{27} (4.15)	3904.6	3905.6	Asp_2 (2.1), Thr_2 (1.8), Ser_1 (0.8), Glu_4 (4.3), Gly_1 (1.0), Ala_3 (3.0), Val_1 (1.2), Ile_1 (1.9), Leu_4 (4.1), Tyr_1 (0.8), Lys_1 (1.1), Arg_3 (3.1), Pro_2 (2.0)
Fragment 4 Cys^{28}-Gly^{65} (4.16)	4882.2	4883.1	Asp_4 (3.8), Thr_2 (1.8), Ser_1 (0.9), Glu_7 (7.4), Gly_2 (2.1), Ala_3 (2.8), Met_1 (1.1), Ile_1 (0.9), Leu_1 (1.0), Tyr_1 (1.0), Phe_1 (1.1), His_1 (1.0), Lys_2 (2.1), Arg_1 (1.1), Pro_1 (1.4)
Fragment 3 Leu^{66}-Gly^{100} (4.17)	3974.29	3973.9	Asp_2 (2.2), Ser_3 (2.7), Glu_5 (5.3), Gly_2 (2.3), Ala_4 (4.1), Val_4 (4.8), Leu_8 (6.5), His_1 (1.2), Lys_1 (1.2), Arg_1 (1.1), Pro_2 (2.1)
Fragment 2 Leu^{101}-Ala^{126} (4.18)	2885.21	2883.9	Asp_1 (1.1), Thr_2 (1.7), Ser_3 (2.0), Glu_2 (2.1), Gly_1 (1.2), Ala_6 (5.9), Ile_1 (1.1), Leu_5 (4.9), Lys_1 (1.7), Arg_2 (2.4), Pro_2 (2.2)
Fragment 1 Ala^{127}-Arg^{166} (4.19)	5157.26	5158.8	Asp_3 (3.0), Thr_3 (3.7), Ser_1 (1.1), Glu_1 (1.0), Gly_3 (2.6), Ala_3 (2.6), Val_1 (1.9), Ile_1 (1.1), Leu_5 (5.4), Tyr_2 (1.5), Phe_3 (3.8), Lys_3 (3.6), Arg_6 (5.6), Pro_1 (1.0)

Table 4.1 Results for Synthesis of Fragments for TAEC

Fragments 2-5 were synthesised on semicarbazide linker resin which was generated immediately prior to use. The final alanine residue of Fragment 5 was coupled as the Tbfmoc amino acid, Tbfmoc acting as the N^{α} protecting group during the ligation steps and facilitating a final affinity purification step of the protein. Fragment 1 was synthesised on Wang resin. All peptides were cleaved under acidiolytic conditions and purified by semi-preparative HPLC.

4.8 Attempted Ligation of Fragment 4 and Fragment 5

The ligation of Fragments 4 and 5 was proposed to proceed in an analogous manner to that of native chemical ligation.

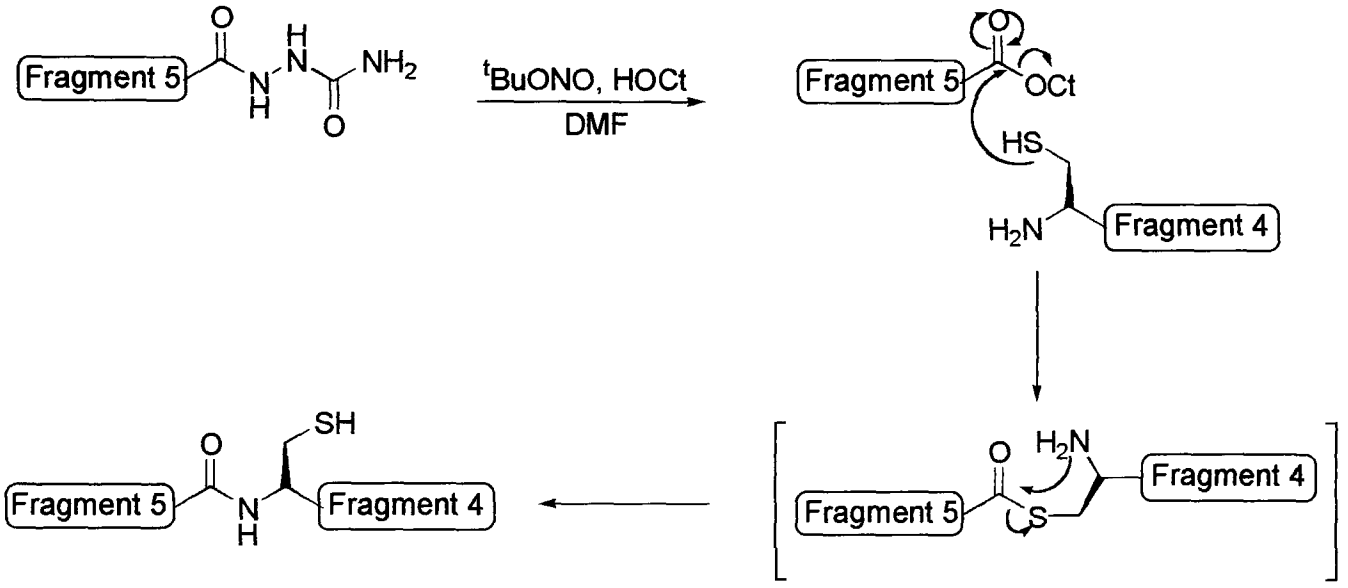


Figure 4.7 Proposed Ligation of Fragments 5 and 4

Conversion of Fragment 5 to an α-thioether derivative, used in native chemical ligation, was deemed not necessary due to the highly reactive nature of the peptide-OCt derivative. Conversion of Fragment 5 to its peptide-Oct derivative was attempted. However, despite the promising results which indicated that conversion of the semicarbazide to the active ester is only marginally slower than that with the corresponding hydrazine,[16] the reaction was slow. Even after 24 hours little change was noted by HPLC analysis. Attempted ligation of Fragment 4 surprisingly led to the disappearance of the 364nm peak, characteristic of the Tbfmoc group, on HPLC,

indicating that the DIEA was causing some form of decomposition or cleavage of the Tbfmoc moiety.

4.8.1 Tbfmoc Cleavage Study

In order to assess the stability of the Tbfmoc group studies were carried out using a variety of bases. Tbfmoc-Leu-OH was stirred in a solution of DMF containing the base under investigation. An aliquot of the reaction mixture was removed hourly and analysed by HPLC to check for the presence of Tbfmoc. The results are summarised in **Table 4.2**.

Base	pKa[20]	Stability
DIEA	10.98[21]	< 1 hour
$NaHCO_3$	10.33	< 2 hours
2,6-Lutidine	6.60	< 4 hours
2,4,6-Collidine	7.43	< 4 hours
Pyridine	5.23	< 8 hours

Table 4.2 Stability of Tbfmoc Group

The stability was taken to be the time after which decomposition of the Tbfmoc group was evident by HPLC. From this it can be seen that pyridine is the most effective base since the Tbfmoc group appears to be stable for up to 8 hours. Accordingly, ligation of Tbfmoc protected fragments should be possible using pyridine. Another complementary approach to overcome this problem is incorporation of the, DIEA stable, Fmoc group for N^{α} protection.

4.9 Synthesis of Fragment 5A employing the Fmoc Group for N^{α} Protection

The synthesis of Fragment 5 was repeated employing the Fmoc group for N^{α} protection. Recent research has demonstrated that potential *bis*-acyl hydrazide formation from peptide hydrazides is only problematic when working at low temperatures.[16] As TAEC avoids the need for low temperatures this competing side reaction should not be observed. Synthesis of the peptide was carried out on a hydrazide linker, in an effort to enhance the formation of the peptide-OCt derivative.

4.9.1 Attempted Ligation of Fragment 5A and Fragment 4

Conversion of the peptide-hydrazide to the active ester was attempted under similar conditions to those outlined in **Section 4.8**. After 6 hours there was inconclusive evidence by HPLC for the formation of the active ester. Nevertheless, ligation was attempted. The peptide component of the reaction was isolated by gel filtration. Subsequent analysis of this material by electrospray mass spectrometry indicated that both fragments were obtained intact from the reaction. As the peptide hydrazide was isolated from the reaction it was assumed that azide formation had not taken place.

4.10 Trial Ligation

Due to the disappointing results obtained in **Section 4.9.1**, ligation of smaller fragments was attempted. If successful, it may be presumed that steric hindrance or secondary structure formation may be causing problems in the formation of the larger peptide azide intermediate.

Two peptide fragments, both five residues in length (modelling the site of the second potential coupling in dhEPO) were synthesised. The *N* terminal fragment was synthesised on the semicarbazide linker resin with the Fmoc group employed for N^{α} protection.

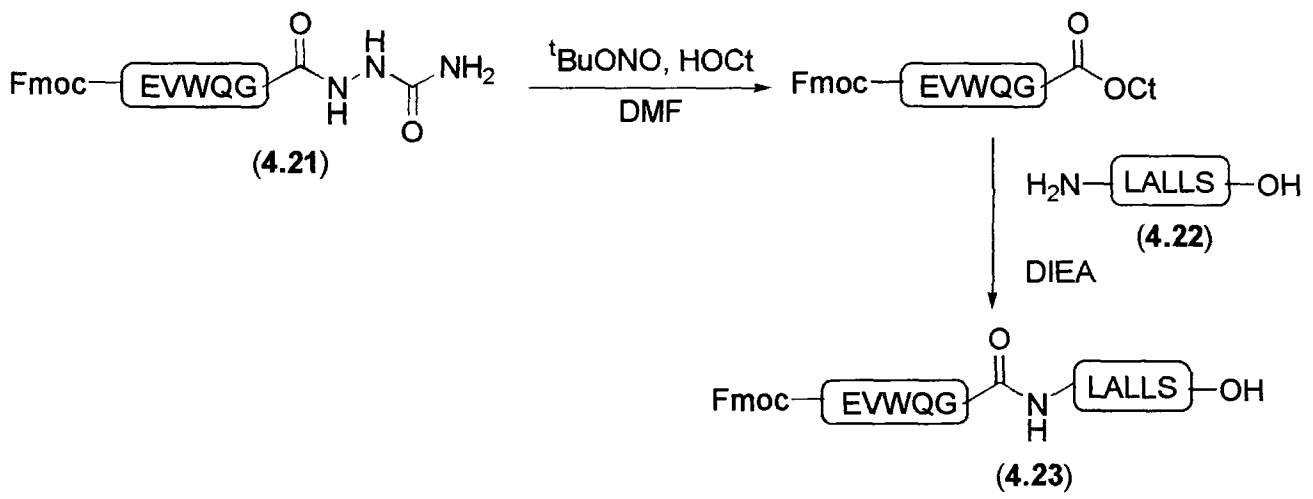


Figure 4.8 Trial Ligation of Small Fragments

Coupling was then attempted using the protocol outlined for the larger fragments.

Formation of the active ester was observed by HPLC. The reaction was deemed complete after 4 hours, although only approximately 80% of the starting material had converted to the active ester.

Fragment 4 was added and the reaction left for 1 hour prior to analysis. The coupled product was clearly visible by HPLC, subsequent analysis of this material confirmed that the coupling had proceeded satisfactorily.

4.11 Conclusions

Coupling of the dhEPO fragments *via* TAEC, using either the hydrazide or semicarbazide peptide derivative, was unsuccessful. Initial studies indicated that the Tbfmoc group was not stable to DIEA, prompting a series of studies into the stability of Tbfmoc to a variety of bases. Of the bases tested pyridine appeared to offer the greatest stability and it is likely that it will prove useful for future work involving ligation of Tbfmoc protected material.

Ligation of two smaller peptide fragments has been seen to proceed swiftly using TAEC methodology, conversion of the semicarbazide peptide to the peptide-OCt derivative posing no particular problems.

Therefore, it appears likely that the problems associated with the larger fragments may be due to steric interactions or the formation of some degree of secondary structure. As a result the *C*-terminal functional group may be buried in the core of the peptide and therefore, inaccessible to reagents. It is possible that the use of structure disrupting agents may help in overcoming such problems. Substituting ^{t}BuONO with *iso*-amyl nitrite, which has been shown to be more effective in the formation of sterically hindered azides,[9] may also prove fruitful.

---

1. M. Schnözler and S. B. H. Kent, *Science*, 1992, **256**, 221.

2. L. E. Canne, S. J. Bark and S. B. H. Kent, *J. Am. Chem. Soc.*, 1996, **118**, 5891.

3. P. E. Dawson, M. J. Churchill. M. R. Ghadiri and S. B. H. Kent, *J. Am. Chem. Soc.*, 1997, **119**, 4325.

4. M. Baca, T. W. Muir, M. Schnözler and S. B. H. Kent, *J. Am. Chem. Soc.*, 1995, **117**, 1881.

5. W. Lu, M. A. Qasim and S. B. H. Kent, *J. Am. Chem. Soc.*, 1996, **118**, 8518.

6. In *"The Peptides, Analysis, Synthesis, Biology"*, Volume I, E. Gross and J. Meienhofer (Eds), Academic Press, Inc., 1979.

7. Y. S. Klausner and M. Bodansky, *Synthesis*, 1974, **8**, 549.

8. T. Curtius, *Ber. Dtsch. Chem. Ger.*, 1902, **35**, 3226.

9. J. Honzel and J. Rudinger, *Coll. Czech. Chem. Commun.*, 1961, **26**, 2333.

10. F. Weygand, A. Prox and W. König, *Chem. Ber.*, 1966, **99**, 1451.

11. G. W. Anderson, J. E. Zimmerman and F. M. Callahan, *J. Am. Chem. Soc.*, 1966, **88**, 1338.

12. P. Sieber, B. Riniker,M. Brugger, B. Kamber and W. Rittel, *Helv. Chim. Acta*, 1970, **53**, 2135.

13. D. S. Kemp, Z. Bernstein and J. Rebek Jr., *J. Am. Chem. Soc.*, 1970, **92**, 4756.

14. P. Wang, R. Layfield, M. Landon, R. J. Mayer and R. Ramage, *Tetrahedron Lett.*, 1998, **39**, 8711.

15. P. Wang, K. T. Shaw, B. Whigham and R. Ramage, *Tetrahedron Lett.*, 1998, **39**, 8719.

16. P. Wang, *Personal Communication*, University of Edinburgh.

17. R. Ramage, S. L. Irving and C. M^{c}Innes, *Tetrahedron Lett.*, 1993, **34**, 6599.

18. J. Patterson, *Personal Communication*, University of Edinburgh.

19. A. R. J. Comer, *Ph.D. Thesis*, University of Edinburgh, 1996.

20. In 'Handbook of Physical Properties of Organic Compounds' (Eds P. W. Howard and W. M. Meylan), CRC Press, Inc. 1997.

21. pKa obtained using the ACD / ilab web service version 2.7 at http://acdlabs.com/ilab/

CHAPTER 5
Experimental

5.1 Notes

All chemicals were purchased from the Aldrich Chemical Company or Fischer Acros. Purity was checked by melting point, if solid, and ^{1}H NMR spectroscopy. All amino acids were purchased from Bachem, NovaBiochem or PerSeptive Biosystems and are of the *L* configuration unless otherwise stated. Tricyclic amide linker resin was purchased from Bachem. *p*-Alkoxybenzylalcohol (Wang) resin was purchased from Bachem or NovaBiochem. Hydrazine resin was purchased from NovaBiochem. Peptide synthesis grade dimethylformamide (DMF), 1,4-dioxane and piperidine were purchased from Rathburn Chemicals. *N,N*-diisopropylethylamine (DIEA) and trifluoroacetic acid (TFA) were peptide synthesis grade and were purchased from Applied Biosystems (ABI). Sephadex G-75 and G-50 were purchased from Pharmacia.

Melting points were recorded in open capillaries using a Buchi 510 oil immersion melting point apparatus and are uncorrected. Analytical thin layer chromatography (tlc) was performed using plastic sheets precoated with silica gel $60F_{254}$ (Merck) in the solvent systems indicated in the text. Compounds were visualised by absorption at 254nm. Optical rotations were measured on an A1000 polarimeter (Optical Activity Ltd) in a 10cm cell in the solvents indicated in the text. Infra red (IR) spectra were recorded on a BioRad SPC3200 FT-IR spectrometer. Ultra violet (UV) spectra were recorded on a Perkin Elmer single beam spectrometer in the solvents indicated in the text. Sonication was carried out at room temperature in a Decon F5300b sonic bath. Elemental analyses were performed on a Perkin Elmer 2400 CHN analyser. Proton NMR spectra were recorded on a Brucker WP-200 (200 MHz) or a Varian Gemini 2000 (200 MHz) spectrometer in the solvents indicated in the text. All chemical shifts were referenced to TMS. All coupling constants (*J*) are quoted in Hz Carbon-13 NMR spectra were recorded on an AC250 (63 MHz)

spectrometer in the solvents indicated in the text, where q denotes CH_3, t denotes CH_2, d denotes CH and s denotes quaternary C.

High and low resolution fast atom bombardment mass spectra (FAB-MS) were measured on a Kratos MS50TC instrument using thioglycerol, 3-nitrobenzyl alcohol or glycerol as a matrix. High and low resolution electron impact mass spectra (EI-MS) were recorded on a Kratos 920MS instrument using the matrices described previously. Matrix assisted laser desorption ionisation (MALDI) time of flight (ToF) mass spectra were recorded on a Perseptive Biosystems Voyager™ Biospectrometry™ workstation using either α-cyano-4-hydrocinnamic acid or 3,5-dimethoxy-4-hydrocinnamic acid as the matrix. Electrospray mass spectra were recorded on a Micromass Platform II Mass Spectrometer.

Amino acid analyses (AAA) were performed using a Pharmacia Biotech Biochrom 20 amino acid analyser or a LKB 4150 alpha amino acid analyser on the hydrosylate obtained after heating the sample in 6N HCl at 110°C in sealed Carius tubes for the times indicated in the text. The HCl was removed on a Savant Speed Vac Plus SC110A after hydrolysis and the residue reconstituted in 0.2M citrate buffer pH 2.2.

High performance liquid chromatography (HPLC) was carried out using either an ABI system comprising 2 x 1406A solvent delivery systems, a 1480 injector/mixer and a 1783 detector/controller or a Gilson system comprising 2 x 306 solvent delivery systems, a 811c dynamic mixer, a 805 manomeric module, a 119 UV/vis detector and a Gilson 715 gradient controller. Compounds were eluted from the various columns described in the text using a linear gradient of acetonitrile (far UV grade, Rathburn chemicals) in Milli-Q water with both solvents containing 0.1% TFA (HPLC grade, Fisons). All experiments were run on a gradient of 10% - 90% MeCN over 30 minutes unless otherwise stated. Columns referred to throughout the text for analytical HPLC are Aquapore C_4 (100 x 4.6mm, 7μm), Aquapore C_8 (220 x 4.6mm, 7μm), Vydac C_8 (250 x 4.6mm, 7μm) and Vydac C_{18} (250 x 4.6mm, 7μm). For semi-preparative HPLC the columns employed were Vydac C_8 (250 x 10mm, 10μm) and Vydac C_{18} (250 x 10mm, 10μm).

5.2 Solid Phase Peptide Synthesis

Automated peptide synthesis was carried out on an ABI 430A synthesiser with on line UV monitoring using an ABI 758A detector. All peptides were synthesised using the 9-fluorenylmethoxycarbonyl (Fmoc) base labile N^{α} protecting group in conjunction with orthogonal acid labile side chain protecting groups and an acid labile peptide-resin linker.

5.2.1 *C*-Terminal Acids

The *C*-terminal Fmoc amino acid (6 equivalents) was dissolved in the minimum amount of DMF and DIC (3 equivalents) added. The solution was sonicated at room temperature for 15 mins before being added to Wang resin (1g) pre-swollen in the minimum volume of DMF with a catalytic amount of DMAP. The mixture was sonicated at room temperature for 3 hours. The resin was collected by filtration, washed with copious quantities of DMF, 1,4-dioxane, DCM, ether and dried.

For a low resin loading (< 0.12 mmol/g) the procedure was modified as follows:
The *C*-terminal Fmoc amino acid (3 equivalents) was dissolved in the minimum amount of DMF and DIC (1.5 equivalents) added. The solution was sonicated at room temperature for 15 mins before being added to Wang resin (1g) pre-swollen in the minimum volume of DMF with a catalytic amount of DMAP. The mixture was left to sit on the bench for 50 minutes at room temperature. The resin was collected by filtration, washed with copious quantities of DMF, 1,4-dioxane, DCM, ether and dried.

5.2.2 *C*-Terminal Semicarbazides

Synthesis of (2-copolystyrene-1%-divinylbenzene)methyl-aminocarbomethoxy-5-(9'-fluorenylmethoxycarbonyl)hydrazine-dibenzocycloheptadiene resin (4.10)[1]

Modification of the resin

Tricyclic amide linker resin (**4.7**) (1g, 0.59 mmol/g) was sonicated in 20% v/v piperidine/DMF for 45 mins. The resin was filtered, washed with DMF, 1,4-dioxane, DCM, ether and dried.

Generation of the isocyanate intermediate

The resin was swollen in the minimum amount of DCM and DIEA (l02μl, 0.59mmol, 1eq) added and the mixture sonicated for 10mins. A solution of triphosgene (0.49g, 1.65mmol, 2.75eq) in DCM (4ml) was added and the mixture sonicated for a further 20mins. The resin was filtered, washed with DCM, ether and dried.

FTIR ν_{max}/cm^{-1} (KBr disc) 2235.9 (N=C=O)

*Trapping of the isocyanate intermediate with 9-Fluorenylmethoxycarbonyl Hydrazine (**4.9**)*

The isocyanate resin (0.9g) was swollen in the minimum volume of DCM. A solution of $FmocNHNH_2$ (**4.9**) (405mg, 1.59mmol, 3eq) in DCM (5ml) was added and the mixture sonicated for 2 hours. The resin was filtered, washed with DMF, 1,4-dioxane, DCM, ether and dried. A KBr disc of the resin was made to ensure complete consumption of the isocyanate species.

5.2.3 Side Chain Protecting Groups

The side chain protecting groups used were as follows: *t*-butyl ethers for Ser, Thr and Tyr, *t*-butyl esters for Asp and Glu, *t*-butoxycarbonyl (Boc) for Lys and Trp, τ-triphenylmethyl (Trt) for Asn, Gln and His, 2,2,5,7,8-pentamethylchroman-6-sulfonyl (Pmc) or 2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl (Pbf) for Arg and Trt or 4-picolyl (Pic) for Cys.

5.2.4 The Fmoc Loading Test

A sample of dry, accurately weighed, Fmoc amino acid resin (3-10mg) was sonicated in 20% piperidine in DMF (10ml) for 10 mins at room temperature. The UV spectrum of the supernatant was recorded between 320 and 280nm. The resin functionality (mmol/g) was then calculated using the Beer-Lambert law (ε302 = 15400 for the fulvene - piperidine adduct).

5.2.5 The Tbfmoc Loading Test

A sample of dry, accurately weighed, Tbfmoc amino acid resin (3-10mg) was sonicated in 20% piperidine in DMF (10ml) for 10 mins at room temperature. The UV spectrum of the supernatant was recorded between 380 and 320nm. The resin functionality (mmol/g) was then calculated using the equation;

$$\text{Resin functionality (mmol/g)} = (0.613 \times \text{Abs}_{364})/\text{weight of resin (mg)}$$

5.2.6 Automated SPPS

Synthetic procedures were programmed into the ABI 430A peptide synthesiser. Each synthetic cycle, resulting in the coupling of a single amino acid, involved the following steps:

1. Capping

   The resin was vortexed with a solution of acetic anhydride (0.5 M), DIEA (0.125 M) and HOBt (0.2% w/v) in DMF/1,4-dioxane (1:1, 10ml) for 10mins. The reaction vessel was then drained and the resin washed 6 times with DMF/1,4-dioxane (1:1, 10ml)

2. Deprotection

   The resin was vortexed with a solution of 20% (v/v) piperidine in DMF/1,4-dioxane (1:1, 10ml) for 10mins. The reaction vessel was drained and an aliquot of the solution passed through the UV detector. The resin was then washed 4 times with DMF/1,4-dioxane (1:1, 10ml), treated with 20% (v/v) piperidine in DMF/1,4-dioxane (1:1, 10ml) for 1.5mins and washed a further 6 times with DMF/1,4-dioxane (1:1, 10ml).

3. Coupling

   Step 1 Activation

   The Fmoc amino acid (1 mmol) was treated with HOCt (1 mmol) and DIC (1 mmol) in DMF/1,4-dioxane (1:1, 8ml) for 15mins to form the active ester which was then transferred to the reaction vessel

   Step 2 Coupling

   The mixture was vortexed for 30mins, the reaction vessel drained and the resin washed 6 times with DMF/1,4-dioxane (1:1, 10ml)

5.3 General Procedure for Cleavage of Peptides from Resin

The resin bound peptide (<1g) was stirred in a scavenger cocktail containing H_2O (1ml), TIS (0.5ml), thioanisole (0.5ml), EDT (0.25ml) and phenol (750mg) for 10 minutes. TFA (10ml) was then added to this and the mixture stirred at room temperature for the times indicated in the text.

5.4 Ellmans Test[2]

0.1 – 0.2mg of the compound was dissolved in 6M guanidine.HCl, 0.1M phosphate buffer, pH 7.3 and 1mmol ethylenediaminetetraacetic acid (EDTA) (2ml). 100µl of a 3mM solution of 5,5'-dithiobis(2-nitrobenzoic acid) (DTNB) in 0.1M phosphate buffer, pH 7.3 was added and the solution mixed thoroughly. The rapid appearance of a bright yellow colour indicated the presence of a free thiol group.

5.5 Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE)

SDS-PAGE experiments were carried out using a Bio-Rad Mini-PROTEAN^{TM} gel electrophoresis apparatus. Proteins were subjected to electrophoresis at 200V constant power at 25°C under denaturing conditions using the Laemmli discontinuous buffer system,[3] and a running buffer prepared with Tris base (9g), glycine (43.2g) and SDS (3g) in 600ml deionized H_2O (this buffer was diluted 5 fold before use). Separating gel (15% acrylamide) was prepared with 1.5M Tris-HCl pH 8.8 (2.5ml), deionized water (2.35ml), 10% w/v SDS/H_2O (100µl) and Bio-Rad's 30% Acrylamide/BIS solutions 37.5:1 mixture (5ml). Stacking gel (4% acrylamide) was prepared with 0.5M Tris-HCl pH 6.8 (2.5ml), deionized water (6.1ml), 10% w/v SDS/H_2O (100µl) and Bio-Rad's 30% Acrylamide/BIS solutions 37.5:1 mixture (1.3ml). Both gels were polymerised with the addition of 10% w/v ammonium persulfate (APS)/H_2O (100µl) and tetramethylethylenediamine (TEMED) (20µl).
Protein samples were prepared by heating the sample in the appropriate amount of denaturing buffer (deionized H_2O (3.8ml), 0.5M Tris-HCl pH 6.8 (1ml), glycerol

(0.8ml), 10% w/v SDS/H_2O (1.6ml), 2-mercaptoethanol (0.4ml) and 1% w/v bromophenol blue/H_2O (0.4ml)) at 90°C for 5 minutes. All samples were detected by staining with 0.1% w/v Coomassie Blue R-250/AcOH:MeOH:H_2O (1:4:5) for 1 hour followed by destaining in AcOH:MeOH:H_2O (1:4:5) until the protein bands could be visualised (generally 2 hours).

5.6 Experimental Details

2-(3-pyridinyl)-1,3-thiazolidine-4-carboxylic acid (2.3)

L-Cysteine (5.76g, 50mmol) and 3-pyridine-carboxaldehyde (4.7ml, 50mmol) were heated under reflux in 50% v/v ethanol/H_2O (100ml) for 1 hour. The solvent was removed under reduced pressure to give a yellow solid which was recrystallised from ethanol, affording *2-(3-pyridinyl)-1,3-thiazolidine-4-carboxylic acid* as a mixture of diastereomers.[4] ^{1}H NMR spectroscopy indicated a 60 : 40 ratio of diastereomers by integration of the epimeric CH. **^{1}H NMR (200MHz, d_6DMSO δH/ppm)** 5.55 (1H, s) = 40%, 5.74 (1H, s) = 60%.

Yield 7.63g, 73%; **Mp** 159-161°C; **CHN** found C 51.33%, H 4.85% N 12.85% ($C_9H_{10}N_2O_2S$ requires C 51.41%, H 4.79% N 13.31%); **MS** (EI) m/z = 210 (M^+); **HRMS** (EI) 210.0457 (M^+ requires 210.0463); **^{1}H NMR (200MHz, d_6DMSO δH/ppm)** (* denotes major diastereomer) 3.10 (1H, dd, J = 9.9, 9,6, βH), *3.13 (1H, dd, J = 10.3, 4.9, βH), 3.35 (1H, dd, J = 9.9, 6.9, β'H), *3.37 (1H, dd, J = 10.3, 6.8, β'H), 3.93 (1H, dd, J = 6.9, 9.9, αH), *4.21 (1H, dd, J = 4.9, 6.8, αH), 5.55 (1H, s, epimeric CH), *5.74 (1H, s, epimeric CH), *7.37 (1H, dt, J = 0.7, 7.8, aromatic H), 7.39 (1H, dt, J = 0.7, 7.8, aromatic H), *7.84 (1H, dt, J = 7.9, 1.7, aromatic H), 7.98 (1H, dt, J = 7.9, 1.9, aromatic CH), *8.47 (1H, dd, J = 4.8, 1.7, aromatic H), 8.52 (1H, dd, J = 4.8, 1.6, aromatic H), *8.62 (1H, d, J = 2.2, aromatic H), 8.69 (1H, d, J = 2.2, aromatic H); **^{13}C {^{1}H} NMR (63MHz, d_6DMSO δC/ppm)** 38.13 (t), 38.34 (t), 64.93 (d), 65.89 (d), 68.54 (d), 69.12 (d), 123.45 (d), 123.64 (d), 134.76 (d), 135.23 (d), 148.40 (d), 148.76 (d), 149.44 (d), 135.08 (s), 137.41 (s), 172.05 (s), 172.87 (s); **λ_{max}/nm (MeOH, ε/dm^{3}mol^{-1}cm^{-1})** 262 (3335); **FTIR ν_{max}/cm^{-1} (bromoform)** 1137

(C-O), 1602 (aromatic), 1713 (C=O), 2888, 2969 (CH), 3017 (aromatic CH), 3262 (NH), 3433 (OH); $[\alpha]_D^{24}$ –32.9° (c = 0.46, DMF).

Attempted Cleavage of the Protecting Group

$NaBH_3CN$ (375mg, 6mmol, 3eq) was added to a suspension of *2-(3-pyridinyl)-1,3-thiazolidine-4-carboxylic acid* (**2.3**) (500mg, 2mmol) in AcOH (5ml) at 0°C. H_2O (20ml) was added to this and the resulting solution allowed to stir at 0°C overnight. [5] The solvent was removed under reduced pressure to yield a yellow oil. This was added to MeCN with stirring, resulting in the precipitation of a white solid, which was filtered, washed and dried.

MS (FAB) 213 (MH^+); **^{1}H NMR (200MHz, D_2O, δH/ppm)** 2.91 (1H, dd, J = 5.2, 14.5, βH), 3.01 (1H, dd, J = 4.9, 14.6, β'H), 3.69 (1H, dd, J = 5.2, 5.0, αH), 4.15 (2H, dd, J = 4.6, 13.3, CH_2), 7.46 (1H, dd, J = 5.0, 8.0, aromatic H), 7.92 (1H, dt, J = 8.0, 1.9, aromatic H), 8.50 (1H, dd, J = 1.6, 5.0, aromatic H), 8.54 (1H, d, J = 2.1, aromatic H); **Ellmans test** +ve

2-(4-pyridinyl)-1,3-thiazolidine-4-carboxylic acid (2.4)

L-Cysteine (5.76g, 50mmol) and 4-pyridine-carboxaldehyde (4.7ml, 50mmol) were heated under reflux in 50% v/v ethanol/H_2O (100ml) for 1 hour. The solvent was removed under reduced pressure to give a yellow solid which was recrystallised from ethanol, affording *2-(4-pyridinyl)-1,3-thiazolidine-4-carboxylic acid* as a mixture of diastereomers.[6] ^{1}H NMR spectroscopy indicated a 60 : 40 ratio of diastereomers by integration of the epimeric CH. **^{1}H NMR (200MHz, d_6DMSO δH/ppm)** 5.54 (1H, s) = 40%, 5.76 (1H, s) = 60%.

Yield 8.54g, 85%; **Mp** 160-163°C; **CHN** C 51.16%, H 4.98%, N 12.91% ($C_9H_{10}N_2O_2S$ requires C 51.41%, H 4.79% N 13.31%); **MS** (FAB) m/z = 211 (MH^+), **HRMS** 211.0537 (MH^+ requires 211.0541); **^{1}H NMR (200MHz, d_6DMSO δH/ppm)** (* denotes major diastereomer) 3.05 (1H, dd, J = 6.8, 8.9, βH), *3.09 (1H, dd, J = 6.1, 8.9, βH), *3.31 (1H, dd, J = 6.8, 8.9, β'H), 3.35 (1H, dd, J = 6.8, 6.9, β'H), 3.97 (1H, dd, J = 6.9, 8.9, αH), *4.10 (1H, dd, J = 6.1, 6.8, αH), 5.54 (1H, s, epimeric CH), *5.76 (1H, s, epimeric CH), *7.25 (2H, dd, J = 6.1, 2.9, 2 aromatic H), 7.40 (2H, dd, J = 6.1, 2.9, 2 aromatic H), 8.45 (2H, dd, J = 6.1, 2.9, 2 aromatic

CH), *8.51 (2H, dd, J = 6.1, 2.9, 2 aromatic CH); **^{13}C {^{1}H} NMR (63MHz, d_6DMSO δC/ppm)** 38.11 (t), 38.30 (t), 65.04 (d), 65.88 (d), 69.27 (d), 69.94 (d), 121.79 (d), 122.37 (d), 148.54 (d), 149.72 (d), 149.88 (d), 148.83 (s), 151.38 (s), 151.50 (s), 172.16 (s), 172.74 (s); **FTIR ν_{max}/cm^{-1} (bromoform)** 1139 (C-O), 1613 (aromatic), 1716 (C=O), 2889, 2930, 2970 (CH), 3017 (aromatic CH), 3315 (NH), 3409 (OH); **λ_{max}/nm (MeOH, ε/dm^{3}mol^{-1}cm^{-1})** 260 (3938); **$[\alpha]_D^{24}$**–34.5° (c = 0.47, DMF).

Attempted Cleavage of the Protecting Group

Activated zinc dust (600mg, 9mmol, 4.5eq) was added to a suspension of *2-(4-pyridinyl)-1,3-thiazolidine-4-carboxylic acid* (**2.4**) (500mg, 2mmol) in 50% v/v AcOH/H_2O (20ml). The resulting solution was left to stir for 2 hours at room temperature. The excess zinc dust was removed by filtration and the solvent was removed under reduced pressure. The resulting white solid was filtered, washed and dried.

MS (FAB) 213 (MH^+); **^{1}H NMR (200MHz, D_2O, δH/ppm)** 2.87 (1H, dd, J = 4.9, 11.3, βH), 2.98 (1H, dd, J = 4.8, 11.3, β'H), 3.67 (1H, dd, J = 4.8, 5.0, αH), 4.08 (2H, s, CH_2), 7.31 (2H, dd, J = 6.2, 1.8, 2 aromatic H), 8.49 (2H, dd, J = 6.1, 1.8, 2 aromatic H); **Ellmans test** +ve.

***S*-3-Picolyl-*L*-Cysteine (2.7)**

3-Picolyl Chloride (HCl salt) (9.02g, 50mmol) was added to a stirred solution of *L*-cysteine (6.05g, 50mmol) in 2N NaOH (50ml) and ethanol (60ml). After 2 hrs a white precipitate formed which was filtered, washed with ethanol, ether and dried *in vacuo* to yield *S-3-picolyl-L-cysteine* as a white solid.

Yield 9.11g, 86%; **Mp**. 208-212°C; **CHN**: found, C 51.31%, H 5.62%, N 13.05% ($C_9H_{12}N_2O_2S$ requires C 50.92%, H 5.70%, N 13.20%); **MS** (FAB) m/z = 213 (MH^+); **HRMS** (FAB) 213.0690 (MH^+ requires 213.0698); **^{1}H NMR (200MHz, D_2O δH/ppm)** 2.77 (1H, d, J = 6.59, βH), 2.78 (1H, d, J = 4.76, β'H) 3.64 (2H, s, picolyl CH_2), 3.70 (1H, dd, J = 4.76, 6.59, αCH), 7.24 (1H, dd, J = 4.76, 7.69, aromatic CH), 7.69 (1H, dt, J = 7.69, 1.83, aromatic CH), 8.23 (1H, dd, J = 1.47, 4.76, aromatic CH), 8.31 (1H, d, J = 1.83, aromatic CH); **^{13}C {^{1}H} NMR (63MHz,**

D_2O δC/ppm) 31.34 (t), 31.21 (t), 53.31 (d), 134.27 (s), 124.28 (d), 138.01 (d), 147.40 (d), 148.68 (d), 172.69 (s); **FTIR ν_{max}/cm^{-1} (bromoform)** 1139 (C-O), 1516 (aromatic), 1596 (C=O), 2118, 2497 (NH_2), 2886, 2968 (CH), 3017 (aromatic CH), 3442 (OH); **λ_{max}/nm (MeOH, $\varepsilon/dm^3mol^{-1}cm^{-1}$)** 264 (3098); **Ellmans test** (-ve); **Crystal Data** $C_9H_{12}N_2O_2S$, $M = 212.27$, monoclinic, $a = 9.934(2)$, $b = 8.777(2)$, $c = 11.122(3)$ Å, $U = 965.7(3)$ Å^3, $T = 220(2)$ K, space group $P21$, $Z = 4$, $\mu = 2.792$ mm^{-1}, 3843 reflections measured, 2328 unique ($R_{int} = 0.0919$) which were used in all calculations. The final $wR(F^2)$ was 0.1359.

_S_-4-Picolyl-L-Cysteine (2.8)

4-Picolyl chloride (HCl salt) (9.02g, 50mmol) was added to a stirred solution of *L*-cysteine (6.05g, 50mmol) in 2N NaOH (50ml) and ethanol (60ml). After 2 hrs a white precipitate formed which was filtered, washed with ethanol, ether and dried *in vacuo* to yield *S-4-picolyl-L-cysteine* as a white solid.

Yield (9.91g, 93%); **Mp**. 207-209°C (Lit.[7] 209.5-211°C); **CHN**: found, C 51.30%, H 5.70%, N 13.37% ($C_9H_{12}N_2O_2S$ requires C 50.92%, H 5.70%, N 13.20%); **MS** (FAB) m/z = 213 (MH^+); **HRMS** (FAB) 213.0697 (MH^+ requires 213.0698); **^{1}H NMR (200MHz, D_2O δH/ppm)** 2.78 (1H, d, J = 4.76, βH), 2.80 (1H, d, J = 6.59, β'H), 3.64 (2H, s, picolyl CH_2), 3.72 (1H, dd, J = 4.76, 6.59, αH), 7.29 (2H, d, J = 6.26, 2 aromatic H), 8.28 (2H, d, J = 6.26, 2 aromatic H); **^{13}C {^{1}H} NMR (63MHz, D_2O δC/ppm)** 31.41 (t), 34.02 (t), 53.31 (d), 124.61 (d), 148.67 (d), 148.65 (s)*, 172.68 (s) * quaternary carbon assigned from low noise decoupled spectrum; **FTIR ν_{max}/cm^{-1} (bromoform)** 1140 (C-O), 1527 (aromatic), 1602 (C=O), 2123, 2577 (NH_2), 2920 (CH), 3017 (aromatic CH), 3421 (OH); **λ_{max}/nm (MeOH, $\varepsilon/dm^3mol^{-1}cm^{-1}$)** 260 (6057); **Ellmans test** (-ve); **Crystal Data** $C_9H_{12}N_2O_2S$, $M = 212.27$, triclinic, $a = 5.7568(10)$, $b = 8.186(2)$, $c = 11.172(2)$ Å, $U = 491.08(15)$ Å^3, $T = 293(2)$ K, space group $P1$, $Z = 2$, $\mu = 2.745$ mm^{-1}, 5316 reflections measured, 3400 unique ($R_{int} = 0.0132$) which were used in all calculations. The final $wR(F^2)$ was 0.0986.

N^{α}-9-Fluorenylmethoxycarbonyl -*S*-3-Picolyl-*L*-Cysteine (2.9)

Fmoc-*N*-hydroxysuccinimide (4.04g, 12mmol) was added to a stirred solution of *S*-3-picolyl-*L*-cysteine (**2.7**) (2.54g, 12mmol) in 10% Na_2CO_3/H_2O (60ml) and 1,4-dioxane (60ml) and allowed to stir at room temperature overnight. The solid *N*-hydroxy succinimide was removed by filtration and the filtrate concentrated to half volume, acidified with AcOH and extracted with EtOAc (3 x 150ml). The organic extracts were combined, washed with brine (1 x 100ml), dried (Na_2SO_4) and the solvent removed under reduced pressure. The resulting solid was purified by wet flash chromatography ($CHCl_3$:MeOH:AcOH; 85:10:5) to yield *N^{α}-9-fluorenylmethoxycarbonyl-S-3-picolyl-L-cysteine* as a white solid.

Yield 3.87g, 74%; **Mp**. 146-148°C; **Tlc** R_f 0.33 ($CHCl_3$:MeOH:AcOH; 85:10:5); **MS** (FAB) m/z = 435 (MH^+); **HRMS** (FAB) 435.1397 (MH^+ requires 435.1379); **^{1}H NMR (200MHz, d_6DMSO δH/ppm)** 2.69 (1H, dd, J = 13.7, 9.5, βH), 2.83 (1H, dd, J = 13.7, 4.6, β'H), 3.80 (2H, s, picolyl CH_2), 4.15 – 4.34 (4H, m, fluorenyl CH, CH_2, αH), 7.29 – 7.44 (5H, m, 4 fluorenyl aromatic H, 1 picolyl aromatic H), 7.74 (2H, d, J = 7.3, fluorenyl aromatic H), 7.81 (1H, d, J = 8.4, picolyl aromatic H), 7.89 (2H, d, J = 7.3, fluorenyl aromatic H), 8.45 (1H, dd, J = 1.6, 4.8, picolyl aromatic H), 8.52 (1H, d, J = 1.8, picolyl aromatic H); **^{13}C {^{1}H} NMR (63MHz, d_6DMSO δC/ppm)** 32.34 (t), 46.77 (d), 53.84 (d), 65.92 (t), 120.30 (d), 123.73 (d), 125.46 (d), 127.26 (d), 136.61 (d), 148.25 (d), 149.96 (d), 123.73 (s), 134.43 (s), 140.89 (s), 143.95 (s), 156.21 (s), 172.39 (s); **FTIR ν_{max}/cm^{-1} (bromoform)** 1140 (C-O), 1507 (aromatic), 1601 (C=O), 1704 (urethane), 2890, 2969 (CH), 3017 (aromatic CH), 3322 (NH), 3408 (OH); **λ_{max}/nm (MeOH, $\varepsilon/dm^3mol^{-1}cm^{-1}$)** 300 (4883), 289 (4557), 265 (18120); **$[\alpha]_D^{24}$** –31.2° (c = 1.1, DMF); **Ellmans test** (-ve); **HPLC** (Aquapore RP18 C18 λ = 214nm) 50% MeCN, R_t 15min.

N^{α}-9-Fluorenylmethoxycarbonyl -*S*-4-Picolyl-*L*-Cysteine (2.10)

Fmoc-*N*-hydroxysuccinimide (4.04g, 12mmol) was added to a stirred solution of *S*-4-picolyl-*L*-cysteine (**2.8**) (2.54g, 12mmol) in 10% Na_2CO_3/H_2O (60ml) and 1,4-dioxane (60ml) and allowed to stir at room temperature overnight. The solid *N*-hydroxy succinimide was removed by filtration and the filtrate concentrated to half volume, acidified with AcOH and extracted with EtOAc (3 x 150ml). The organic

extracts were combined, washed with brine (1 x 100ml), dried (Na_2SO_4) and the solvent removed under reduced pressure. The resulting solid was purified by wet flash chromatography ($CHCl_3$:MeOH:AcOH; 85:10:5) to yield *N^{α}-9-fluorenylmethoxycarbonyl-S-4-picolyl-L-cysteine* as a white solid.

Yield 3.48g, 67%; **Tlc** R_f 0.33 ($CHCl_3$:MeOH:AcOH; 85:10:5); **Mp**. 142-145°C; **MS** (FAB) m/z = 435 (MH^+); **HRMS** (FAB) 435.1371 (MH^+ requires 435.1379); **^{1}H NMR (200MHz, d_6DMSO δH/ppm)** 2.72 (1H, dd, J = 9.3, 13.7, βH), 2.86 (1H, dd, J = 4.6, 13.7, β'H), 3.77 (2H, s, picolyl CH_2), 4.14 – 4.34 (4H, m, fluorenyl CH,CH_2, αH), 7.28 – 7.43 (6H, m, fluorenyl aromatic H), 7.69 – 7.75 (2H, m, fluorenyl aromatic H), 7.87 (2H, d, J = 7.4, picolyl aromatic H), 8.49 (2H, d, J = 5.5, picolyl aromatic H); **^{13}C {^{1}H} NMR (63MHz, d_6DMSO δC/ppm)** 32.54 (t), 34.27 (t), 46.78 (d), 53.85 (d), 65.93 (t), 120.27 (d), 125.03 (d), 125.44 (d), 127.25 (d), 127.83 (d), 148.15 (d), 140.89 (s), 143.94 (s), 150.05 (s), 156.21 (s), 172.28 (s); **FTIR ν_{max}/cm^{-1} (bromoform)** 1139 (C-O), 1530 (aromatic), 1611 (C=O), 1686 (urethane), 2890, 2970 (CH), 3017 (aromatic CH), 3350 (NH), 3466 (OH); **λ_{max}/nm (MeOH, ε/dm^{3}mol^{-1}cm^{-1})** 300 (4991), 288 (5859), 265 (17143); **$[\alpha]_D^{24}$** –34.0° (c = 0.1, DMF); **Ellmans test** (-ve); **HPLC** (Aquapore RP18 C18, λ = 214nm) 50% MeCN, R_t 15 min.

Synthesis of H_2N-Tyr-Phe-Cys-His-Leu-Asp-Ile-Ile-Trp-OH incorporating Fmoc-*S*-3-picolyl-*L*-cysteine (2.11)

The synthesis was carried out on a 0.25mmol starting from Fmoc-Trp(Boc) functionalised Wang resin (500mg, 0.51mmol/g). All amino acids were single coupled as the active HOCt ester. On completion of the synthesis the resin (730mg) was subjected to acidiolytic cleavage for 1 hour. The resin was removed by filtration and the peptide precipitated into ice cold ether, centrifuged to a pellet, washed with ether (3 x 50ml), dissolved in MeCN/H_2O and lyophilised.

Yield 218mg; **MS** MALDI ToF 1301.29 (MH^+, MW 1300.42); **HPLC** (Hichrom C_8/C_{18}) R_t 20.2mins 63% MeCN; **Amino Acid Analysis** (24 hour hydrolysis) Asp_1 (1.08), Ile_2 (1.91), Leu_1 (1.09), Tyr_1 (0.79), Phe_1 (0.98), His_1(1.14).

Attempted Cleavage of the picolyl protecting group

Peptide (**2.11**) (50mg, 0.04mmol) was dissolved in 20% v/v AcOH/H_2O (25ml) and $NaBH_3CN$ (7.5mg, 0.12mmol, 3eq) added. The solution was left to stir at room temperature and samples removed for analysis by HPLC at hourly intervals. After 6 hours there was no change by HPLC and the reaction was left overnight. Analysis by HPLC confirmed that only starting material was present.

Synthesis of H_2N-Tyr-Phe-Cys-His-Leu-Asp-Ile-Ile-Trp-OH incorporating Fmoc-*S*-4-picolyl-*L*-cysteine (2.12)

The synthesis was carried out on a 0.25mmol starting from Fmoc-Trp(Boc) functionalised Wang resin (400mg, 0.66mmol/g). All amino acids were single coupled as the active HOCt ester. On completion of the synthesis the resin (635mg) was subjected to acidiolytic cleavage for 1 hour. The resin was removed by filtration and the peptide precipitated into ice cold ether, centrifuged to a pellet, washed with ether (3 x 50ml), dissolved in MeCN/H_2O and lyophilised.

Yield 205mg; **MS** MALDI ToF 1301.54 (MH^+, MW 1300.42); **HPLC** (Vydac C_{18}) R_t 16.4mins, 53% MeCN; **Amino Acid Analysis** (24 hour hydrolysis) Asp_1 (1.14), Ile_2 (1.46), Leu_1 (1.22), Tyr_1 (0.96), Phe_1 (1.01), His_1 (1.29).

Attempted Cleavage of the picolyl protecting group

Peptide (**2.12**) (50mg, 0.04mmol) was dissolved in 50% v/v AcOH/H_2O (10ml). Activated zinc dust (500mg, 7.7mmol, 200eq) was added and the solution allowed to stir at room temperature. Samples were removed at hourly intervals and analysed by HPLC. After 2 hours all starting material had disappeared and the reaction was deemed complete. Co-injection with a sample of the peptide prepared with Cys(Trt) confirmed that cleavage of the protecting group had taken place.

Yield 36mg; **MS** MALDI ToF 1210.36 (MH^+, MW 1209.43); **HPLC** (Vydac C_{18}) R_t 17.8mins, 56% MeCN; **Amino Acid Analysis** (24 hour hydrolysis) Asp_1 (1.12), Cys_1 (0.89), Ile_2 (1.78), Leu_1 (1.36), Tyr_1 (1.05), Phe_1 (1.09), His_1 (1.21).

Synthesis of H_2N-Ala-Cys(Pic)-Gly-OH (2.13)

The synthesis was carried out on a 0.25mmol scale starting from Fmoc-Gly functionalised Wang resin (400mg, 0.65mmol/g). Both residues were single coupled as the active HOCt ester. The resin was subjected to acidiolytic cleavage for 2 hours, the peptide precipitated into ice cold ether, centrifuged to a pellet and washed with ether (3 x 50ml). The peptide was dissolved in 50% v/v MeCN/H_2O and lyophilysed.

^{1}H NMR (250MHz, CD_3OD δH/ppm) 1.61 (3H, d, J = 7.1, CH_3), 2.84 (1H, dd, J = 7.8, 14.0, βH Cys), 3.07 (1H, dd, J = 6.2, 14.0, β'H Cys), 3.93 – 4.13 (5H, m, picolyl CH_2, αH Ala, CH_2 Gly), 4.72 (1H, dd, J = 6.2, 7.8, αH Cys), 8.04 (2H, d, J = 6.6, aromatic CH), 8.80 (2H, d, J = 6.6, aromatic CH); **HPLC** (Vydac C_{18}) R_t 21.8mins, 66% MeCN; **MS** MALDI-ToF 379.27 (K^+ salt, $C_{14}H_{20}N_4O_4S$ requires 340.4).

OXYTOCIN

Synthesis of Oxytocin with 2 Picolyl Protecting Groups (2.14)

The synthesis was carried out on a 0.25mmol scale, starting from Fmoc-Gly functionalised Wang resin (600mg, 0.44mmol/g). All amino acids were single coupled as their active HOCt ester. On completion of the synthesis the resin (820mg) was subjected to acidiolytic cleavage for 2 hours. The resin was removed by filtration and the peptide precipitated into ice cold ether, centrifuged to a pellet and washed a further 3 times with ether. The pellet was dissolved in MeCN:H_2O (1:1) and lyophilised (**Yield** 190mg). 50 mg of this material was purified by semi preparative HPLC.

Yield 22mg; **HPLC** (Vydac C_{18}) R_t 16 mins, 53% MeCN; **MS** ES-MS 1192.71 (M^+), MALDI-ToF 1193.88 (MH^+) $C_{55}H_{77}N_{13}O_{13}S_2$ MW 1192.41; **Amino Acid Analysis** (24 hours hydrolysis) Asp_1 (1.03), Glu_1 (1.08), Gly_1 (1.03), Ile_1 (0.97), Leu_1 (1.04), Tyr_1 (0.83), Pro_1 (1.03); **Ellmans Test** – ve.

Cleavage of the picolyl groups (2.15)

Oxytocin(Pic) (**2.14**) (10mg) was dissolved in a solution of 50% v/v AcOH/H_2O (8 ml) and activated zinc dust (500 mg) added to this and the solution left to stir at room temperature. After 30 mins an aliquot was removed and analysed by HPLC (Vydac

C_{18}), which indicated that cleavage of the picolyl groups was complete. The excess zinc dust was removed by filtration and the material desalted by semi preparative HPLC.

Yield 8.2mg; **HPLC** (Vydac C_{18}) R_t 18.4 mins, 59% MeCN; **MS** ES-MS 1010.8 (M^+), $C_{43}H_{67}N_{11}O_{13}S_2$ MW 1010.19; **Amino Acid Analysis** (24 hours hydrolysis) Asp_1 (0.87), Glu_1 (1.07), Gly_1 (1.28), Cys_2 (0.68), Ile_1 (1.03), Leu_1 (1.39), Tyr_1 (1.08), Pro_1 (1.53); **Ellmans Test** + ve.

Disulfide Formation (2.16)

Oxytocin (**2.15**) (8mg) was added to a solution of 5% v/v DMSO/TFA (5ml) and left to stir at room temperature. Samples were removed hourly and analysed by HPLC (Vydac C_{18}). After 4 hours the reaction was deemed complete. The TFA and DMSO were removed *in vacuo* and the resulting yellow oil added drop-wise to ice cold ether. The precipitated peptide was centrifuged to a pellet and lyophilised from MeCN/H_2O.

Yield 6.8mg; **HPLC** (Vydac C_{18} λ = 214nm) R_t 16.8 mins, 55% MeCN; **MS** ES-MS 1008.6 (M^+), $C_{43}H_{65}N_{11}O_{13}S_2$ MW 1008.17; **Amino Acid Analysis** (24 hours hydrolysis) Asp_1 (0.97), Glu_1 (1.04), Gly_1 (1.01), Cys_2 (0.34), Ile_1 (0.96), Leu_1 (1.03), Tyr_1 (0.98), Pro_1 (1.00); **Ellmans Test** – ve.

ENDOTHELIN-1

Synthesis of Endothelin-1 with 2 picolyl protected cysteine

The synthesis was carried out on a 0.25mmol scale starting from Fmoc-Trp(Boc) functionalised resin (500mg, 0.51mmol/g). All amino acids were single coupled as the active HOCt ester. On completion of the synthesis the resin was collected by filtration, washed with DMF, 1,4-dioxane, DCM, ether and dried.

Yield 1.2g; **Final Fmoc loading test** 0.07mmol/g.

Loading Tbfmoc to the resin

Fmoc-ET-1-resin (700mg) was sonicated in 20% v/v piperidine/DMF for 20 minutes to remove the Fmoc protecting group. The resin was filtered, washed with DMF,

1.4-dioxane, DCM, ether and dried. The resin was swollen in the minimum volume of dry DCM, Tbfmoc chloroformate (69mg, 0.15mmol, 3eq) and DIEA (8.5µl, 0.05mmol, 1eq) added and the mixture sonicated in the dark for 3 hours. The resulting Tbfmoc-ET-1-resin was collected by filtration washed with DCM, ether and dried.

Amino acid analysis (24 hours hydrolysis) Asp_2 (2.09), Ser_3 (2.24), Glu_1 (1.42), Cys_2 (1.24), Val_1 (1.58), Met_1 (0.92), Ile_2 (2.34), Leu_2 (3.40), Tyr_1 (0.93), Phe_1 (1.42), His_1 (1.21), Lys_1 (0.21).

Cleavage of Tbfmoc-ET-1 from the resin

Tbfmoc-ET-1-resin (700mg) was added to H_2O (1ml), thioanisole (0.5ml), TIS (0.5ml), EDT (0.25ml), phenol (750mg) and allowed to stir under an atmosphere of nitrogen for 20 minutes. TFA (10ml) was then added and the mixture left to stir under nitrogen in the dark for 1 hour. The resin was removed by filtration and the peptide precipitated into ice cold ether, centrifuged to a pellet, washed with ether (3 x 50ml), dissolved in 50% v/v MeCN/H_2O and lyophilised.

Yield 300mg; **Amino acid analysis** (24 hours hydrolysis) Asp_2 (2.14), Ser_3 (1.98), Glu_1 (1.45), Cys_2 (1.03), Val_1 (1.61), Met_1 (0.94), Ile_2 (2.28), Leu_2 (2.73), Tyr_1 (1.05), Phe_1 (1.45), His_1 (1.27), Lys_1 (1.09).

Affinity Purification of Tbfmoc-ET-1-OH on Charcoal

Tbfmoc-ET-1-OH (60mg) was dissolved in 7M urea (25ml). Freshly washed charcoal (800mg) was added and the mixture vortexed for 10 minutes. The charcoal was centrifuged (3500rpm, 10mins) to a pellet. HPLC analysis of the supernatant confirmed all the Tbfmoc tagged material had bound to the charcoal. The charcoal pellet was washed with 50% v/v 7M urea/IPA (50ml, 5 times) until no absorbance at 214nm was detectable by HPLC. The peptide was then cleaved from the charcoal by treatment with 10% v/v piperidine in 7M urea:IPA (1:1, 20ml), the charcoal was removed by centrifugation (3500rpm, 10mins), the pH of the supernatant adjusted to pH 4.5 with AcOH. The IPA was removed *in vacuo* and the material desalted on a Sephadex G50 (medium grade) gel filtration column using 20% v/v AcOH/H_2O as eluent. Peptide containing fractions were pooled and lyophilised.

Yield 14.7mg; **MS** ES-MS 2678.3 (MW 2677.89); **HPLC** (Vydac C_{18}) R_t 20.8 mins, 65% MeCN; **Amino acid analysis** (24 hours hydrolysis) Asp_2 (2.08), Ser_3 (2.31), Glu_1 (1.33), Cys_2 (1.31), Val_1 (1.26), Met_1 (0.89), Ile_2 (1.77), Leu_2 (2.14), Tyr_1 (1.07), Phe_1 (1.15), His_1 (1.09), Lys_1 (1.11).

Disulfide formation in ET-1

ET-1 (4.5mg) was dissolved in 20% v/v AcOH/H_2O (4ml) and left to stir for a few minutes. DMSO (1ml) was added to the solution to give a final concentration of 20% DMSO. The solution was left to stir at room temperature and samples removed hourly for analysis by HPLC. After 3 hours the reaction appeared to be complete, the solvent was removed under reduced pressure and the resulting oil added dropwise to ether to precipitate the peptide. This was then isolated by centrifugation, washed with ether (3 x 20ml), dissolved in 50% v/v MeCN/H_2O and lyophilised.

Yield 3.2mg; **HPLC** (Vydac C_{18}) R_t 19.6 mins, 62% MeCN.

Cleavage of the picolyl protecting groups

ET-1 (2mg) was added to 50% v/v AcOH/H_2O and left to stir until all the peptide had dissolved. Freshly activated zinc dust (200mg) was added to this and the solution left to stir at room temperature. A sample was removed after 1 hour and analysed by HPLC, indicating that a mixture of products had been formed.

Reaction of *L*-Cystine with activated zinc dust

L-cystine (20mg) was dissolved in 50% v/v AcOH/H_2O, activated zinc dust (200mg) added and the solution allowed to stir at room temperature. After 1 hour a sample was removed and analysed with Ellmans reagent. The solution turned bright yellow indicating cleavage of the disulfide bond had occurred.

Synthesis of Endothelin-1 with 4 Picolyl Protecting Groups (2.17)

The synthesis was carried out on a 0.25mmol scale, starting from Fmoc-Trp(Boc) functionalised Wang resin (600mg, 0.44mmol/g). All amino acids were single coupled as their active HOCt ester. The resin was collected by filtration, washed with DMF, 1,4-dioxane, DCM, ether and dried.

Yield 1.4g; **Final Fmoc loading test** 0.1mmol/g.

Loading Tbfmoc to the resin

Fmoc-ET-1(Pic)-resin (1.4g) was sonicated in 20% v/v piperidine/DMF for 20 minutes to remove the Fmoc protecting group. The resin was filtered, washed with DMF, 1.4-dioxane, DCM, ether and dried. The resin was swollen in the minimum volume of dry DCM, Tbfmoc chloroformate (191mg, 0.42mmol, 3eq) and DIEA (24μl, 0.14mmol, 1eq) added and the mixture sonicated in the dark for 3 hours. The resulting Tbfmoc-ET-1-resin was collected by filtration washed with DCM, ether and dried.

Amino Acid Analysis (24 hours hydrolysis) Asp_2 (2.03), Ser_3 (1.79), Glu_1 (1.38), Val_1 (1.28), Met_1 (0.75), Ile_2 (1.50), Leu_2 (1.98), Tyr_1 (1.30), Phe_1 (1.45), His_1 (1.31), Lys_1 (1.24).

Cleavage of Tbfmoc-ET-1(Pic) from the resin

Tbfmoc-ET-1-resin (700mg) was added to H_2O (1ml), thioanisole (0.5ml), TIS (0.5ml), EDT (0.25ml), phenol (750mg) and allowed to stir under an atmosphere of nitrogen for 20 minutes. TFA (10ml) was then added and the mixture left to stir under nitrogen in the dark for 1 hour. The resin was removed by filtration and the peptide precipitated into ice cold ether, centrifuged to a pellet, washed with ether (3 x 50ml), dissolved in 50% v/v MeCN/H_2O and lyophilised.

Yield 404mg; **Amino acid analysis** (24 hours hydrolysis) Asp_2 (2.01), Ser_3 (2.08), Glu_1 (1.21), Val_1 (1.28), Met_1 (0.92), Ile_2 (1.71), Leu_2 (2.38), Tyr_1 (1.05), Phe_1 (1.13), His_1 (1.12), Lys_1 (1.09).

Affinity Purification of Tbfmoc-ET-1(Pic)-OH on Polystyrene

Tbfmoc-ET-1(Pic)-OH (195mg) was dissolved in 40% v/v AcOH/H_2O (45ml). Polystyrene (3.7g) was added and the mixture vortexed for 10 minutes. The polystyrene was removed by filtration. HPLC analysis of the filtrate confirmed all the Tbfmoc tagged material had bound to the polystyrene. The polystyrene was washed with 40% v/v AcOH/H_2O (50ml, 5 times) until no absorbance at 214nm was detectable by HPLC. The peptide was then cleaved from the polystyrene by

treatment with 10% v/v piperidine in 60% v/v MeCN/H_2O (50ml). The solution was acidified to pH 4.5 with AcOH, the MeCN removed *in vacuo* and the material desalted on a Sephadex G50 (medium grade) gel filtration column using 20% v/v AcOH/H_2O as eluent. Peptide containing fractions were pooled and lyophilised. A final purification step by semi-preparative HPLC was required.

Yield 14.7mg; **HPLC** (Vydac C_{18} λ = 214nm) R_t 20 mins, 64% MeCN; **MS** ES-MS 2859.9, MALDI-ToF 2861.87 (MW 2859.89); **Amino Acid Analysis** (24 hours hydrolysis) Asp_2 (2.04), Ser_3 (2.20), Glu_1 (1.25), Val_1 (1.25), Met_1 (0.78), Ile_2 (1.88), Leu_2 (2.20), Tyr_1 (1.10), Phe_1 (1.10), His_1 (1.10), Lys_1 (0.94); **Ellmans Test** –ve.

Cleavage of the picolyl groups (2.18)

Endothelin-1(Pic) (**2.17**) (10mg) was dissolved in a solution of 50% v/v AcOH/H_2O (8 ml) and activated zinc dust (500 mg) added to this and the solution left to stir at room temperature. HPLC analysis (Vydac C_{18}) indicated that cleavage of the picolyl groups was complete after 1 hour. The excess zinc dust was removed by filtration and the material desalted by semi preparative HPLC.

HPLC (Vydac C_{18}) R_t 21.4 mins, 68% MeCN; **MS** ES-MS 2495.1, MALDI-ToF 2497.31 (MW 2495.89); **Amino Acid Analysis** (24 hours hydrolysis) Asp_2 (1.77), Ser_3 (1.27), Glu_1 (1.18), Cys_4 (0.52), Val_1 (1.18), Met_1 (0.69), Ile_2 (2.13), Leu_2 (2.10), Tyr_1 (0.69), Phe_1 (1.12), His_1 (1.02), Lys_1 (1.21); **Ellmans Test** + ve.

Disulfide Formation (2.19)

Endothelin-1 (**2.18**) (5mg) was added to a solution of 5% v/v DMSO/TFA (5ml) and left to stir at room temperature. Samples were removed hourly and analysed by HPLC (Vydac C_{18}). After 6 hours the reaction was complete, giving two products of which the major was identified as the correct material form injection with a commercial standard (Peptide Institute INC., Japan). The TFA and DMSO were removed *in vacuo* and the resulting yellow oil was added drop-wise to ice cold ether. The precipitated peptide was centrifuged to a pellet dissolved in MeCN/H_2O and purified by semi-preparative HPLC.

HPLC (Vydac C_{18}) R_t 19mins, 60% MeCN; **MS** ES-MS 2491.22 (MW 2491.89); **Amino Acid Analysis** (24 hours hydrolysis) Asp_2 (2.02), Ser_3 (1.96), Glu_1 (1.19),

Cys_4 (1.29), Val_1 (1.04), Met_1 (0.97), Ile_2 (1.91), Leu_2 (2.07), Tyr_1 (0.64), Phe_1 (0.97), His_1 (0.98), Lys_1 (0.96); **Ellmans Test** –ve.

Chemical Synthesis and Purification of Deglycosylated Human Erythropoietin

HPLC Conditions All HPLC analysis of dhEPO and dhEPO(Pic) was carried out using an Aquapore C_4 column (dimensions 100 x 4.6mm, 7μm). Compounds were eluted from the column using a linear gradient of 10% - 90% MeCN in Milli-Q water over 30 minutes. All experiments were monitored at 214nm and also at 364nm for Tbfmoc material.

Synthesis of dhEPO

The synthesis was carried out on a 0.1mmol scale using Fmoc-Arg(Pmc) functionalised Wang resin (580mg, 0.17mmol/g). Side chain protecting groups were as described in the text. All amino acids were single coupled as the active HOCt ester, with the exception of the last 20 residues which were double coupled. Approximately one third of the resin was removed after Ala^{111} and the synthesis continued to the end. The resin was washed with DMF, dioxane, DCM, ether and dried.

Final Fmoc loading test 0.02mmol/g; **Yield** 1.5g.

Loading the Tbfmoc Group

Fmoc-dhEPO-resin (500mg) was sonicated in 20% v/v piperidine/DMF for 20 minutes to remove the Fmoc protecting group. The resin was filtered, washed with DMF, 1.4-dioxane, DCM, ether and dried. The resin was swollen in the minimum volume of dry DCM, Tbfmoc chloroformate (13.75mg, 0.03mmol, 3eq) and DIEA (1.79μl, 0.01mmol, 1eq) added and the mixture sonicated in the dark for 3 hours. The resulting Tbfmoc-dhEPO-resin was collected by filtration washed with DCM, ether and dried.

Amino Acid Analysis (24 hours hydrolysis) Asx_{12} (12.2), Thr_{11} (11.4), Ser_{10} (8.9), Glx_{19} (21.9), Gly_9 (11.7), Ala_{19} (19.9), Cys_4 (1.5), Val_{11} (10.5), Met_1 (0.8), Ile_5 (4.6), Leu_{23} (21.2), Tyr_4 (4.2), Phe_4 (4.3), His_2 (2.8), Lys_8 (8.7), Arg_{13} (13.5), Pro_8 (9.4);

(36 hours hydrolysis) Asx_{12} (12.0), Thr_{11} (10.8), Ser_{10} (8.5), Glx_{19} (21.0), Gly_9 (11.7), Ala_{19} (19.1), Cys_4 (1.2), Val_{11} (10.0), Met_1 (0.9), Ile_5 (4.6), Leu_{23} (18.7), Tyr_4 (4.8), Phe_4 (4.5), His_2 (2.8), Lys_8 (8.6), Arg_{13} (13.1), Pro_8 (10.4); (48 hours hydrolysis) Asx_{12} (11.9), Thr_{11} (10.9), Ser_{10} (8.4), Glx_{19} (20.9), Gly_9 (11.6), Ala_{19} (19.5), Cys_4 (1.1), Val_{11} (10.4), Met_1 (0.7), Ile_5 (4.7), Leu_{23} (18.8), Tyr_4 (4.9), Phe_4 (4.4), His_2 (2.9), Lys_8 (8.7), Arg_{13} (13.3), Pro_8 (9.9).

Trial Cleavage of Tbfmoc-dhEPO-resin

Tbfmoc-dhEPO-resin (50mg) was added to H_2O (1ml), thioanisole (0.5ml), TIS (0.5ml), EDT (0.25ml), phenol (750mg) and allowed to stir under an atmosphere of nitrogen for 20 minutes. TFA (10ml) was then added and the mixture left to stir under nitrogen in the dark. Samples were removed at hourly intervals and analysed by HPLC. After 4 hours no further change was noted in the HPLC trace and the cleavage was deemed complete.

Cleavage of Tbfmoc-dhEPO-resin

Tbfmoc-dhEPO-resin (500mg) was subjected to the acidiolytic cleavage conditions described above for 4 hours. The resin was separated by filtration and washed with a little TFA (~2ml). Addition of the filtrate to ice-cold ether resulted in the precipitation of the protein. This was centrifuged (3500rpm, 2mins) to a pellet and the supernatant discarded. The pellet was washed a further three times with ether, solubilised in 50% v/v MeCN/H_2O and lyophilised.

Yield 247.9mg; **Amino Acid Analysis** (24 hours hydrolysis) Asx_{12} (12.31), Thr_{11} (10.25), Ser_{10} (7.79), Glx_{19} (21.67), Gly_9 (10.79), Ala_{19} (18.68), Cys_4 (1.98), Val_{11} (13.30), Met_1 (1.19), Ile_5 (5.05), Leu_{23} (25.50), Tyr_4 (2.09), Phe_4 (4.63), His_2 (1.50), Lys_8 (8.99), Arg_{13} (12.41), Pro_8 (8.18).

Affinity Purification of Tbfmoc-dhEPO-OH on Charcoal

Tbfmoc-dhEPO-OH (140mg) was dissolved in 7M urea (25ml). Freshly washed charcoal (1.5g) was added and the mixture vortexed for 10 minutes. The charcoal was centrifuged (3500rpm, 10mins) to a pellet. HPLC analysis of the supernatant confirmed all the Tbfmoc tagged material had bound to the charcoal. The charcoal

pellet was washed with 50% v/v 7M urea/IPA (50ml, 5 times) until no absorbance at 214nm was detectable by HPLC. The protein was then cleaved from the charcoal by treatment with 10% v/v piperidine in 7M urea:IPA (1:1, 20ml), the charcoal was removed by centrifugation (3500rpm, 10mins), the pH of the supernatant adjusted to pH 4.5 with AcOH. The IPA was removed *in vacuo* and the material desalted on a Sephadex G50 (medium grade) gel filtration column using 30% v/v AcOH/H_2O as eluent.

Amino Acid Analysis (24 hours hydrolysis) Asx_{12} (12.07), Thr_{11} (10.38), Ser_{10} (8.24), Glx_{19} (21.38), Gly_9 (11.21), Ala_{19} (18.30), Cys_4 (2.09), Val_{11} (11.30), Met_1 (0.25), Ile_5 (4.60), Leu_{23} (23.68), Tyr_4 (3.89), Phe_4 (4.11), His_2 (2.69), Lys_8 (8.50), Arg_{13} (12.52), Pro_8 (7.77).

Synthesis of dhEPO using the Picolyl Group for Cysteine Protection

The synthesis was carried out on a 0.1mmol scale using Fmoc-Arg(Pbf) functionalised Wang resin (1g, 0.097mmol/g). Side chain protecting groups were as described in the text, all four cysteine residues were protected with the picolyl group. All amino acids were single coupled as the active HOCt ester. Residues Tyr^{156}-Lys^{152}, Val^{144}, Arg^{143}, Lys^{140}, Asp^{136}-Ile^{132} and Pro^{121}, which were found to be low points from a previous synthesis, were coupled using the double cartridge method. Approximately half of the resin was removed after Ala^{98} and the synthesis continued to Pro^{2}. The resin was washed with DMF, dioxane, DCM, ether and dried.

Final Fmoc loading test 0.02mmol/g; **Yield** 1.66g.

N^{α}-(17-Tetrabenzo[*a,c,g,i*]fluorenylmethoxycarbonyl)-*L*-alanine (3.1)

L-Alanine (1.11g, 12.5mmol) was suspended in acetone (125 ml)/water (25ml). Tbfmoc-Cl (4.6g, 10 mmol) and *N,N*-dimethylaniline (3.6 ml, 28.75 mmol, 2.35 eq) were added and the mixture stirred in darkness at room temperature overnight. The solvent was removed under reduced pressure and the resulting orange residue dissolved in DCM (125 ml). This was washed with saturated citric acid (3 x 50 ml), H_2O (6 x 50 ml), dried ($MgSO_4$) and the solvent removed under reduced pressure. The resulting oil was dissolved in DCM and added drop wise to a large volume of hexane, with stirring, to yield an off white solid which was filtered, dried and

purified by wet flash chromatography ($CHCl_3$:MeOH:AcOH; 97:2:1), to yield *N^{α}-(17-Tetrabenzo[a,c,g,i]fluorenylmethoxycarbonyl)-L-alanine* as an off white solid. **Yield** 897mg, 17.5%; **Mp** 133-135°C; **R_f** = 0.35 ($CHCl_3$:MeOH:AcOH; 97:2:1); **MS** (FAB) 512 (MH^+); **HRMS** (FAB) 512.1861 (MH^+ requires 512.1862); **^{1}H NMR (250MHz, $CDCl_3$ δH/ppm)** 1.33 (3H, d, J = 7.1, CH_3), 4.22 – 4.35 (2H, m, αCH, NH), 4.73 (1H, dd, J = 5.2, 11.1, Tbfmoc CH), 5.03 – 5.19 (2H, m, CH_2), 7.55 (8H, m, aromatic CH), 8.22 – 8.33 (2H, m, aromatic CH), 8.58 – 8.78 (6H, m, aromatic CH); **^{13}C {^{1}H} NMR (63MHz, $CDCl_3$ δC/ppm)** 17.98 (q), 47.53 (d), 49.24 (d), 69.15 (t), 123.09 (d), 123.19 (d), 123.48 (d), 124.32 (d), 125.01 (d), 125.09 (d), 125.52 (d), 125.92 (d), 126.06 (d), 126.91 (d), 127.41 (d), 127.89 (d), 128.68 (s), 130.19 (s), 130.30 (s), 131.49 (s), 136.59 (s), 136.92 (s), 141.15 (s), 142.57 (s), 155.72 (s), 176.99 (s); **FTIR ν_{max}/cm^{-1} (bromoform)** 1141 (C-O), 1507 (aromatic), 1711 (C=O & urethane), 3017 (aromatic CH), 3419 (OH); **λ_{max}/nm (MeOH, ε/dm^{3}mol^{-1}cm^{-1})** 378 (17374), 363 (18396), 299 (43946), 287 (36281), 276 (32193), 258 (66430), 251 975628), 237 (56721); **$[\alpha]_D^{24}$** –19.9° (c = 1, DMF); **HPLC** (Aquapore C_4, 10% - 40% B, 5 mins, 40% - 90% B, 41 mins, λ = 364nm) 59% MeCN, R_t 18.4 mins.

Coupling of Tbfmoc-Ala-OH

DhEPO(Pic) resin (1.66g, 0.33mmol) was sonicated in 20% v/v piperidine/DMF for 20 minutes to remove the *N^{α}* protecting group. The resin was washed with DMF, dioxane, DCM, ether and dried.

The resin was swollen in the minimum volume of DCM and Tbfmoc-Ala-OH (**3.1**) (68.7mg, 0.13mmol, 4eq), HOBt (18mg, 0.13mmol, 4eq) and DIC (21μl, 0.13mmol, 4eq) added. The resin was then sonicated in the dark at room temperature for 90 minutes. The resin was collected by filtration, washed with DCM, ether and dried.

Amino Acid Analysis (24 hours hydrolysis) Asx_{12} (13.26), Thr_{11} (9.97), Ser_{10} (6.31), Glx_{19} (21.44), Gly_9 (10.74), Ala_{19} (20.96), Val_{11} (10.10), Met_1 (0.63), Ile_5 (6.61), Leu_{23} (19.14), Tyr_4 (2.12), Phe_4 (4.98), His_2 (2.38), Lys_8 (8.88), Arg_{13} (12.76), Pro_8 (8.69).

Cleavage of Tbfmoc-dhEPO(Pic) Resin

Resin bound Tbfmoc-dhEPO(Pic) (500mg) was stirred in TFA (10ml) containing H_2O (1ml), thioanisole (0.5ml), triisopropylsilane (0.5ml), EDT (0.25ml) and phenol (750mg) in the absence of light. After 4 hours the resin was removed by filtration and the filtrate added to ice cold ether. The resulting precipitate was centrifuged to a pellet, washed with ether (3 x 50ml) and the crude protein dissolved in MeCN/H_2O and lyophilised.

Yield 268mg; **Amino Acid Analysis** (24 hours hydrolysis) Asx_{12} (12.82), Thr_{11} (9.31), Ser_{10} (7.17), Glx_{19} (20.27), Gly_9 (9.85), Ala_{19} (19.54), Val_{11} (11.10), Met_1 (1.74), Ile_5 (6.98), Leu_{23} (22.61), Tyr_4 (3.62), Phe_4 (4.15), His_2 (2.11), Lys_8 (8.06), Arg_{13} (11.29), Pro_8 (8.34).

Tbfmoc/Polystyrene Affinity Purification of dhEPO(Pic)

Tbfmoc-dhEPO(Pic) (96mg) was dissolved in 40% v/v AcOH/H_2O (40ml). Polystyrene (1.8g) was added to this and the solution vortexed for 10 minutes and the polystyrene centrifuged to a pellet. HPLC analysis of the supernatant confirmed that all the Tbfmoc tagged material had bound to the polystyrene. The polystyrene pellet was washed with 40% v/v AcOH/H_2O (3 x 50ml) and H_2O (2 x 50ml). The protein was then cleaved from the Tbfmoc-polystyrene by treatment with a solution of 10% v/v piperidine in 60% v/v MeCN/H_2O. The MeCN was removed in vacuo and the material desalted on a Sephadex G50 (medium grade) gel filtration column using 20% v/v AcOH/H_2O as eluent.

Yield 5mg; **Amino Acid Analysis** (24 hours hydrolysis) Asx_{12} (12.37), Thr_{11} (7.89), Ser_{10} (4.73), Glx_{19} (20.48), Gly_9 (9.34), Ala_{19} (20.79), Val_{11} (11.92), Met_1 (1.19), Ile_5 (6.42), Leu_{23} (23.66), Tyr_4 (3.75), Phe_4 (4.56), His_2 (2.06), Lys_8 (8.82), Arg_{13} (12.64), Pro_8 (8.39).

Removal of Tbfmoc and Cleavage of dhEPO(Pic) from the Resin

Resin bound Tbfmoc-dhEPO(Pic) (200mg) was sonicated in a solution of 20% v/v piperidine/DMF (10ml) for 20 minutes. The resin was then collected by filtration, washed with DMF, dioxane, DCM, ether and dried. Cleavage of the protein from the resin and side chain protecting groups was accomplished by stirring the resin in TFA

(10ml) containing H_2O (1ml), thioanisole (0.5ml), triisopropylsilane (0.5ml), EDT (0.25ml) and phenol (750mg) in the absence of light for 4 hours. The resin was removed by filtration and the filtrate added to ice cold ether. The resulting precipitate was centrifuged to a pellet, washed with ether (3 x 50ml), dissolved in 50% v/v MeCN/H_2O and lyophilised to yield dhEPO(Pic) as a white fluffy solid.

Yield 90.5mg; **Amino Acid Analysis** (24 hours hydrolysis) Asx_{12} (12.51), Thr_{11} (9.36), Ser_{10} (7.34), Glx_{19} (21.60), Gly_9 (9.14), Ala_{19} (19.85), Val_{11} (11.42), Met_1 (1.58), Ile_5 (6.63), Leu_{23} (22.13), Tyr_4 (3.74), Phe_4 (3.92), His_2 (2.05), Lys_8 (8.06), Arg_{13} (11.46), Pro_8 (8.19).

Sephadex G75 Gel Filtration of dhEPO(Pic)

Column Conditions Gel filtration was carried out on a Sephadex G75 (superfine grade) column (diameter = 1.5cm, length = 90cm) pre-equilibrated with 20% v/v AcOH/H_2O. A solution of dhEPO(Pic) (10mg) in 6M urea, 0.2M NaCl, 0.1M phosphate buffer pH 7.5 (400µl) was applied to the column. The protein was eluted from the column using 20% v/v AcOH/H_2O (10ml/h) and fractions collected at hourly intervals. All fractions were analysed by UV (280nm) and a sample of all protein containing fractions (1ml) removed and lyophilised. Analysis of this material by SDS-PAGE identified the fractions containing dhEPO, these were pooled and lyophilised. The above procedure was repeated to generate larger quantities of material. Generally three applications of material to the column was required for purification.

1st Pass Down Column Amino Acid Analysis (24 hours hydrolysis) Asx_{12} (12.08), Thr_{11} (9.15), Ser_{10} (7.30), Glx_{19} (19.20), Gly_9 (9.88), Ala_{19} (20.93), Val_{11} (13.21), Met_1 (1.42), Ile_5 (5.88), Leu_{23} (22.31), Tyr_4 (3.02), Phe_4 (4.44), His_2 (3.20), Lys_8 (8.24), Arg_{13} (10.71), Pro_8 (8.12).

2nd Pass Down Column Amino Acid Analysis (24 hours hydrolysis) Asx_{12} (11.83), Thr_{11} (8.94), Ser_{10} (7.87), Glx_{19} (19.56), Gly_9 (9.18), Ala_{19} (18.86), Val_{11} (12.20), Met_1 (2.75), Ile_5 (7.79), Leu_{23} (21.98), Tyr_4 (2.14), Phe_4 (4.66), His_2 (3.87), Lys_8 (7.96), Arg_{13} (11.41), Pro_8 (7.23).

3rd Pass Down Column Amino Acid Analysis (24 hours hydrolysis) Asx_{12} (11.84), Thr_{11} (9.71), Ser_{10} (8.57), Glx_{19} (19.76), Gly_9 (9.18), Ala_{19} (18.94), Val_{11} (11.42),

Met_1 (1.19), Ile_5 (5.23), Leu_{23} (22.91), Tyr_4 (3.63), Phe_4 (4.01), His_2 (2.16), Lys_8 (7.91), Arg_{13} (11.88), Pro_8 (8.19); ***N*-Terminal Sequencing** cycle 1 – Ala 47.4*p*mol; cycle 2 Pro 44.4*p*mol; cycle 3 Pro 27.7*p*mol; cycle 4 not determined; cycle 5 Leu 14.2*p*mol.

Tryptic Digest of dhEPO(Pic)

DhEPO(Pic) (0.5mg) was dissolved in 6M urea, 0.2M phosphate buffer, pH 8 (200µl) and trypsin (5% w/w) added. The solution was mixed and incubated at 37°C. After 2 hours 100µl of the solution was removed and the digest stopped with the addition of 6N HCl (5µl). The solution was separated using analytical HPLC (Vydac C_{18}, 2ml loop, 1ml/min, 10-90% MeCN, 60mins) and the samples analysed by MALDI-ToF MS. After 5 hours the remaining solution was treated in the same manner. The HPLC trace was identical indicating that the digest was complete after 2 hours.

m/z (MALDI-ToF) 797.49 (MH^+ Leu^{5}-Arg^{10}, requires 797.36) 738.96 (MH^+ Tyr^{15}-Lys^{20}, requires 736.42) 2868.6 (MH^+ Glu^{21}-Lys^{45}, requires 2871.2) 969.61 (MK^+ Val^{46}-Lys^{52}, requires 927.47) 2548.76 (MNa^+ Met^{54}-Arg^{76}, requires 2526.34) 2359.52 (MH^+ Gly^{77}-Lys^{97}, requires 2359.2) 602.57 (MH^+ Ala^{98}-Arg^{103}, requires 602.36) 803.16 (MH^+ Ser^{104}-Arg^{110}, requires 803.49) 588.99 (MH^+ Ala^{111}-Lys^{116}, requires 587.35) 1465.21 (MH^+ Glu^{117}-Arg^{131}, requires 1465.76) 923.07 (MH^+ Thr^{132}-Arg^{139}, requires 924.48) 922.99 (MNa^+ Val^{144}-Arg^{150}, requires 898.45) 1005.16 MH^+ Leu^{155}-Arg^{162}, requires 1003.42).

Isoelectric Focusing of dhEPO(Pic)

Ampholytes (pH 3 – 10, Fluka, 1ml) were added to a solution of dhEPO(Pic) (2mg) in 6M urea (17ml) and the solution added to the focusing chamber of the Rotofor® cell. The system was allowed to reach thermal equilibrium at the cooling temperature of 5°C for 10 minutes. Constant power of 15W was applied to the system and the increase in voltage noted over time. After 2 hours the voltage stabilised and the run was left for a further 30 minutes before harvesting. The pH of the resulting fractions was checked and each was analysed for protein content by HPLC. All protein was found in fractions 19 and 20, indicating a pI of 10.0.

Fraction	PH	Fraction	pH	Fraction	pH	Fraction	pH
1	3.5	**6**	5.5	**11**	7.5	**16**	8.5
2	4.0	**7**	6.0	**12**	7.5	**17**	9.0
3	4.5	**8**	7.0	**13**	8.0	**18**	9.5
4	4.5	**9**	X	**14**	X	**19**	10.0
5	5.0	**10**	X	**15**	8.0	**20**	10.0

Molecular Weight Determination of dhEPO(Pic) by FPLC

Column conditions Analysis was carried out using a Superdex^{TM} 75 HR 10/30 column (column volume V_t 44.93cm^3) pre-equilibrated with 6M urea, 0.2M NaCl, 0.1M phosphate buffer pH 7.5 at a flow rate of 0.5ml/min.

The column void volume was calculated using blue dextran (MWt 2,000,000, V_o 8.51ml). Calibration of the column was achieved using a series of molecular weight standards and the elution volume (V_e) of each measured with monitoring at 280nm, ovalbumin (MWt 43,000, V_e 8.63ml), chymotrypsinogen A (MWt 25,000, V_e 8.77ml), ribonuclease A (MWt 13,700, V_e 9.09ml). The elution volume for dhEPO(Pic) was 8.94ml.

Cleavage of the Picolyl Groups from dhEPO(Pic)

A sample of dhEPO(Pic) (3mg) was dissolved in 50% v/v AcOH/H_2O and activated zinc dust (300mg) added. The resulting solution was left to stir for 2 hours at room temperature. The excess zinc dust was then removed by filtration and the material separated by analytical HPLC.

Yield 1.8mg; **Ellmans Test** +ve; **Amino Acid Analysis** (24 hours hydrolysis) Asx_{12} (12.31), Thr_{11} (9.47), Ser_{10} (8.40), Glx_{19} (21.33), Gly_9 (9.64), Ala_{19} (21.71), Cys_4 (0.77), Val_{11} (12.05), Met_1 (1.71), Ile_5 (5.80), Leu_{23} (23.84), Tyr_4 (3.00), Phe_4 (4.18), His_2 (2.26), Lys_8 (7.89), Arg_{13} (11.67), Pro_8 (8.43).

Transfer Active Ester Condensation of dhEPO

9-Fluorenylmethoxycarbonyl Hydrazine (4.9)

A solution of 9-fluorenylmethylchloroformate (1g, 3.8mmol) in MeCN (130 ml), was added dropwise to hydrazine monohydrate (6 ml, 0.12mol) with stirring over a period of 30 mins. The resulting turbid solution was then allowed to stir for a further

30 mins at room temperature in the absence of light. The solvent was removed under reduced pressure to yield *9-fluorenylmethoxycarbonyl hydrazine* as a white solid which was washed with copious quantities of ethanol and dried *in vacuo*.

Yield (930mg, 95%); **Mp** 172-174°C (lit 172-174°C)[8]; **MS** (FAB) m/z 255 (MH^+); **HRMS** 255.1134 (MH^+ requires 255.1134); **CHN** found, C 70.76%, H 5.38%, N 11.08% ($C_{15}H_{14}N_2O_2$ requires C 70.85%, H 5.55%, N 11.02%); **^{1}H NMR (250MHz, d_6DMSO δH/ppm)** 4.08 (2H, br, NH_2), 4.25 (3H, m, fluorenyl CH and CH_2), 7.35 (4H, m, aromatic CH), 7.67 (2H, d, J = 7.6, aromatic CH), 7.87 (2H, d, J = 7.6, aromatic CH), 8.35 (1H, br, NH); **^{13}C {^{1}H} NMR (63MHz, d_6DMSO δC/ppm)** 46.77 (d), 65.73 (t), 120.12 (d), 120.20 (d), 121.47 (d), 125.33 (d), 127.17 (d), 127.38 (d), 127.73 (d), 129.01 (d), 137.52 (s), 139.51 (s), 140.79 (s), 143.91 (s), 158.27 (s); **FTIR ν_{max}/cm^{-1} (bromoform)** 1141 (C-O), 1512 (aromatic), 1694 (C=O), 2899, 2956 (CH), 3018 (aromatic CH), 3207, 3315 (NH); **λ_{max}/nm (MeOH, ε/dm^{3}mol^{-1}cm^{-1})** 300 (5186), 289 (5186), 265 (16510).

3-Hydroxymethyl-3-nitro-1,5-dioxaspiro-5:5-undecane (4.11)

Cyclohexanone (35ml; 0.34mol) and *tris*(hydroxymethyl)nitromethane (51g; 0.34mol) were refluxed in sodium dried benzene (250ml) with a catalytic amount of *p*-TSA using a Dean and Stark trap for 2.5 hours. The reaction was allowed to cool to room temperature, the organic layer washed with H_2O (2 x 100ml), dried ($MgSO_4$) and the solvent removed under reduced pressure to yield *3-hydroxymethyl-3-nitro-1,5-dioxaspiro-5:5 undecane* as an off white solid which was recrystallised from EtOAc / hexane.

Yield 42g, 54%; **Mp** 95-97°C (lit 97-98°C)[9]; **MS** FAB m/z = 232 (MH^+); **HRMS** 232.1189 (MH^+ requires 232.1185); **CHN** found C 52.03%, H 7.41%, N 5.97% ($C_{10}H_{17}NO_5$ requires C 51.94%, H 7.41%, N 6.06%); **^{1}H NMR (250MHz, $CDCl_3$ δH/ppm)** 1.37-1.75 (10H, m, aliphatic ring CH_2), 2.57 (1H, br, OH), 4.02 (2H, s, $\underline{CH_2}OH$), 4.03 (2H, d, J = 12.66, axial dioxan ring 2 x CH), 4.38 (2H, d, J = 12.66, equatorial dioxan ring 2 x CH); **^{13}C {^{1}H} NMR (63MHz, $CDCl_3$ δC/ppm)** 22.25 (t), 22.37 (t), 25.19 (t), 31.66 (t), 32.19 (t), 60.51 (t), 63.38 (t), 86.65 (s), 99.52 (s); **FTIR ν_{max}/cm^{-1} (bromoform)** 1112 (C-O), 1346, 1443 (NO_2), 1544 (C-NO_2), 2856, 2936 (CH), 3417 (OH); **λ_{max}/nm (MeOH, ε/dm^{3}mol^{-1}cm^{-1})** 271 (770).

Pentafluorophenol carbonate of 3-hydroxymethyl-3-nitro-1,5-dioxaspiro-5:5 undecane (4.12)

3-Hydroxymethyl-3-nitro-1,5-dioxaspiro-5:5 undecane (**4.11**) (11.55g; 50mol), triphosgene (5g; 16.6mol) and DIEA (8.7ml; 50mol) were stirred in sodium dried toluene under an inert atmosphere of dry N_2. After 15 mins a white precipitate formed (DIEA.HCl) and the first stage of the reaction was deemed complete. The chloroformate was then reacted immediately *in situ* by the addition of a mixture of pentafluorophenol (10.1g; 55mol) and DIEA (8.7ml; 50mol) dissolved in the minimum volume of toluene. The reaction was stirred for a further 15 mins then rapidly washed with iced water (2 x 100ml), brine (100ml), dried ($MgSO_4$) and the solvent removed under reduced pressure to yield the *pentafluorophenol carbonate of 3-hydroxymethyl-3-nitro-1,5-dioxaspiro-5:5 undecane* as a white solid which was recrystallised from EtOAc / hexane.

Yield 12.35g; 56%; **Mp** 119-122°C (lit 121-123°C)[10]; **MS** (FAB) m/z = 442 (MH^+); **HRMS** (FAB) 442.0925 (MH^+ requires 442.0925); **CHN** found C 46.39%, H 3.76%, N 3.19% ($C_{17}H_{16}F_5NO_7$ requires C 46.27%, H 3.65%, N 3.17%); **^{1}H NMR (250MHz, $CDCl_3$ δH/ppm)** 1.24 – 1.84 (10H, m, aliphatic ring CH_2), 4.09 (2H, J = 12.66, axial dioxane ring 2 x CH), 4.43 (2H, d, J = 12.66, equatorial dioxane ring 2 x CH), 4.49 (2H, s, CH_2O); **^{13}C {^{1}H} NMR (63MHz, $CDCl_3$ δC/ppm)** 22.23 (t), 22.35 (t), 25.14 (t), 29.79 (t), 33.86 (t), 60.41 (t), 68.26 (t), 83.22 (s), 99.99 (s), 126.17 – 128.34 (aromatic CF), 129.32 (s), 150.45 (s); **^{19}F {^{1}H} NMR (250MHz, $CDCl_3$ δF/ppm)** 10.95, 11.26, 11.78 (CF); **FTIR ν_{max}/cm^{-1} (bromoform)** 1140 (C-O), 1392, 1448 (NO_2), 1521 (aromatic), 1547 (C-NO_2), 1786 (C=O), 2873, 2940, 2957 (CH), 3379 (OH); **λ_{max}/nm (MeOH, ε/dm^{3}mol^{-1}cm^{-1})** 263 (1208), 243 (735).

N^{α}-9-Fluorenylmethoxycarbonyl-N^{ε}-1,5–dioxaspiro–5:5–undecane–3–nitro–3–methoxycarbonyl-lysine (4.13)

Fmoc-Lys-OH (5g; 13.5mmol) was suspended in 1,4-dioxane:H_2O (2:1, 150ml) and triethylamine (3.75ml; 27mmol) added. The solution was cooled to 0°C and the pentafluorophenol carbonate of 3-hydroxymethyl-3-nitro-1,5-dioxaspiro-5:5 undecane (**4.12**) (6.17g; 14mmol) added. The temperature was allowed to rise and the reaction stirred at room temperature overnight. The solution was filtered to

remove any insoluble material and the solvent removed *in vacuo*. The resulting residue was dissolved in EtOAc (100ml) and washed with brine (3 x 50ml), dried ($MgSO_4$) and the solvent removed under reduced pressure. The resulting yellow foam was dissolved in DCM and purified by wet flash chromatography (DCM:MeOH; 9:1) to yield *N^{α}-9-fluorenylmethoxycarbonyl-N^{ε}-1,5–dioxaspiro–5:5–undecane–3–nitro–3–methoxycarbonyl-lysine* as a white solid.

Yield 7.23g, 85%; **Tlc** R_f 0.31 (DCM:MeOH; 9:1); **Mp** 176–177°C (lit 172°C)[10]; **MS** (FAB) m/z = 626 (MH^+); **HRMS** (FAB) 626.2694 (MH^+ requires 626.2714); **^{1}H NMR (200MHz, δH/ppm)** 1.16 – 1.98 (18H, m, aliphatic ring CH_2, lysine side chain CH_2),4.1 – 4.4 (10H, m, dioxane ring CH_2, αH, Fmoc CH, Fmoc CH_2, CH_2O), 7.2 – 7.4 (4H, m, aromatic CH), 7.7 (2H, d, J = 7.3, aromatic CH), 7.9 (2H, d, J = 7.3, aromatic CH); **^{13}C {^{1}H} NMR (63MHz, δC/ppm)** 22.2 (t), 22.6 (t), 22.8 (t), 25.0 (t), 28.3 (t), 29.3 (t), 31.1 (t), 31.9 (t), 46.9 (d), 55.2 (d), 60.5 (t), 63.2 (t), 65.4 (t), 85.6 (s), 98.7 (s), 120.2 (d), 125.4 (d), 127.2 (d), 127.7 (d), 140.6 (s), 144.1 (s), 155.0 (s), 155.8 (s), 176.1 (s); **FTIR ν_{max}/cm^{-1} (bromoform)** 1140 (C-O), 1341, 1446 (NO_2), 1515 (aromatic), 1549 (C-NO_2), 1709 (C=O), 2859, 2937 (CH), 3020 (aromatic CH), 3331 (NH), 3416 (OH); **λ_{max}/nm (MeOH, ε/dm^{3}mol^{-1}cm^{-1})** 300 (7422), 288 (10156), 265 (23438); **$[\alpha]_D^{24}$** –5.8° (c 1.0, DMF); **HPLC** (Aquapore RP18 C18, λ = 214nm) 66% MeCN, R_t 21.2min.

Synthesis of Fragment 1 (4.19)

The synthesis was carried out on a 0.1mmol scale, starting from Fmoc-Arg(Pbf) functionalised Wang resin (1g, 0.1mmol/g). All amino acids were single coupled as their active HOCt ester. On completion of the synthesis the resin (600mg) was subjected to acidiolytic cleavage for 4 hours. The resin was removed by filtration and the peptide precipitated into ice cold ether, centrifuged to a pellet and washed a further 3 times with ether. The pellet was then lyophilised from MeCN:H_2O (**Yield** 152mg). Purification of this material was achieved by semi-preparative HPLC.

Yield 14.2mg; **HPLC** (Vydac C_8) R_t 21.2 mins, 67% MeCN; **MS** ES-MS 5158.8, (MW 5157.26); **Amino Acid Analysis** (24 hours hydrolysis) Asp_3 (3.01), Thr_5 (3.73), Ser_1 (1.07), Glu_1 (0.99), Gly_3 (2.62), Ala_3 (2.60), Val_1 (1.87), Ile_1 (1.06), Leu_5 (5.46), Tyr_2 (1.51), Phe_3 (3.83), Lys_3 (3.68), Arg_6 (5.56), Pro_1 (1.01).

Synthesis of Fragment 2 (4.18)

The synthesis was carried out on a 0.1mmol scale, starting from semi-carbazide linker resin (600mg, 0.19mmol/g). All amino acids were single coupled as their active HOCt ester. On completion of the synthesis the resin (450mg) was subjected to acidiolytic cleavage for 4 hours. The resin was removed by filtration and the peptide precipitated into ice cold ether, centrifuged to a pellet and washed a further 3 times with ether. The pellet was then lyophilised from MeCN:H_2O (**Yield** 120mg). This material was purified by semi-preparative HPLC

Yield 22mg; **HPLC** (Vydac C_{18}) R_t 14.8 mins, 50% MeCN; **MS** ES-MS 2883.9, MALDI-ToF 2889.68 (MW 2885.21); **Amino Acid Analysis** (24 hours hydrolysis) Asp_1 (1.07), Thr_2 (1.67), Ser_3 (2.00), Glu_2 (2.12), Gly_1 (1.18), Ala_6 (5.90), Ile_1 (1.07), Leu_5 (4.93), Lys_1 (1.68), Arg_2 (2.36), Pro_2 (2.17).

Synthesis of Fragment 3 (4.17)

The synthesis was carried out on a 0.1mmol scale, starting from semi-carbazide linker resin (600mg, 0.23mmol/g). All amino acids were single coupled as their active HOCt ester. On completion of the synthesis the resin (500mg) was subjected to acidiolytic cleavage for 4 hours. The resin was removed by filtration and the peptide precipitated into ice cold ether, centrifuged to a pellet and washed a further 3 times with ether. The pellet was then lyophilised from MeCN:H_2O (**Yield** 140mg). This material was purified by semi-preparative HPLC.

Yield 55mg; **HPLC** (Vydac C_{18}) R_t.17.4 mins, 56% MeCN; **MS** ES-MS 3959.69 (-NH), MALDI-ToF 3973.72, (MW 3974.29), **Amino Acid Analysis** (24 hours hydrolysis) Asp_2 (2.24), Ser_4 (2.74), Glu_5 (5.31), Gly_2 (2.25), Ala_4 (4.14), Val_4 (4.84), Leu_8 (6.45), His_1 (1.23), Lys_1 (1.22), Arg_1 (1.07), Pro_2 (2.10).

Synthesis of Fragment 4 (4.16)

The synthesis was carried out on a 0.1mmol scale, starting from semi-carbazide linker resin (600mg, 0.23mmol/g). All amino acids were single coupled as their active HOCt ester. On completion of the synthesis the resin (700mg) was subjected to acidiolytic cleavage for 4 hours. The resin was removed by filtration and the peptide precipitated into ice cold ether, centrifuged to a pellet and washed a further 3

times with ether. The pellet was then lyophilised from MeCN:H_2O (**Yield** 230mg). This material was purified by semi-preparative HPLC.

Yield 78mg; **HPLC** (Vydac C_8) R_t 20 mins, 64% MeCN; **MS** ES-MS 4883.1, (MW 4882.2); **Amino Acid Analysis** (24 hours hydrolysis) Asp_4 (3.79), Thr_2 (1.79), Ser_1 (0.87), Glu_7 (7.38), Gly_2 (2.12), Ala_3 (2.81), Val_5 (5.09), Met_1 (1.12), Ile_1 (0.94), Leu_1 (0.96), Tyr_1 (0.99), Phe_1 (1.11), His_1 (0.98), Lys_2 (2.10), Arg_1 (1.07), Pro_1 (1.38).

Synthesis of Fragment 5 (4.15)

The synthesis was carried out on a 0.1mmol scale, starting from semi-carbazide linker resin (1g, 0.11mmol/g). All amino acids were single coupled as their active HOCt ester. The synthesis was stopped after coupling of Pro^2 to enable manual coupling of Tbfmoc-Ala-OH. The resin (1.2g, 0.084mmol/g) was swollen in dry DCM and Tbfmoc-Ala-OH (174mg, 0.34mmol, 4eq), HOCt (53.4mg, 0.34mmol, 4eq), DIC (53µl, 0.34mmol, 4eq) added. The mixture was sonicated in the dark for 3 hours after which time the resin was separated by filtration, washed with DCM, ether and dried. On completion of the synthesis the resin (600mg) was subjected to acidiolytic cleavage for 4 hours. The resin was removed by filtration and the peptide precipitated into ice cold ether, centrifuged to a pellet and washed a further 3 times with ether. The pellet was then lyophilised from MeCN:H_2O (**Yield** 170mg). This material was purified by semi-preparative HPLC.

Yield 63mg; **HPLC** (Vydac C_8) R_t 25.2 mins, 77% MeCN; **MS** ES-MS 3905.6, (MW 3904.6); **Amino Acid Analysis** (24 hours hydrolysis) Asp_2 (2.10), Thr_2 (1.79), Ser_1 (0.76), Glu_4 (4.29), Gly_1 (1.02), Ala_3 (2.95), Val_1 (1.22), Ile_2 (1.89), Leu_4 (4.07), Tyr_1 (0.84), Lys_1 (1.11), Arg_3 (3.06), Pro_2 (2.00).

Attempted Ligation of Fragment 5 and Fragment 4

Fragment 5 (**4.15**) (10mg, 2.56µmol) and HOCt (157mg, 1mmol) were added to dry DMF (1ml) and left to stir under an atmosphere of dry Argon. ^{t}BuONO (3.04µl, 25.6µmol, 10eq) was added and the solution left to stir at room temperature overnight. Analysis by HPLC indicated the formation of 2 new products. Fragment 4 (**4.16**) (13.7mg, 2.82µmol, 1.1eq) and DIEA (250µl, 1.45mmol) were dissolved in

dry DMF (400µl) and added to the reaction. After one hour a sample of the reaction was analysed by HPLC, indicating the disappearance of all 364nm active material.

Tbfmoc Cleavage Studies

General Procedure Tbfmoc-Leu-OH (2mg, 3.6µmol) and HOCt (157mg, 1mmol) were dissolved in dry DMF (1ml) under an atmosphere of dry Argon. Base (1.45mmol) was added to this and the solution left to stir at room temperature. Samples were removed hourly and analysed by HPLC at 364nm to check for the presence of the Tbfmoc group.

Base	Stability of Tbfmoc group
DIEA	< 1 hour
$NaHCO_3$	< 2 hours
2,6-Lutidiene	< 4 hours
2,4,6-Collidine	< 4 hours
Pyridine	< 8 hours

Synthesis of fragment 5 using the Fmoc group for N^{α} protection (4.20)

The synthesis was carried out on a 0.1mmol scale, starting from chlorotrityl hydrazide linker resin (700mg, 0.13mmol/g). All amino acids were single coupled as their active HOCt ester. On completion of the synthesis the resin (900mg) was subjected to acidiolytic cleavage for 4 hours. The resin was removed by filtration and the peptide precipitated into ice cold ether, centrifuged to a pellet and washed a further 3 times with ether. The pellet was then lyophilised from MeCN:H_2O (**Yield** 180mg). This material was purified by semi-preparative HPLC.

Yield 82mg; **HPLC** (Vydac C_{18}) R_t 24.2 mins, 74% MeCN; **MS** ES-MS 3362.9, (MW 3362); **Amino Acid Analysis** (24 hours hydrolysis) Asp_2 (2.01), Thr_2 (1.76), Ser_1 (0.88), Glu_4 (4.29), Gly_1 (1.14), Ala_3 (2.93), Val_1 (1.26), Ile_2 (1.92), Leu_4 (3.98), Tyr_1 (0.78), Lys_1 (1.19), Arg_3 (3.01), Pro_2 (1.83).

Attempted Ligation of Fragment 5 and Fragment 4

Fragment 5 (**4.20**) (5mg, 1.37µmol) and HOCt (157mg, 1mmol) were added to dry DMF (1ml) and left to stir under an atmosphere of dry Argon. ^{t}BuONO (1.63µl, 13.7µmol, 10eq) was added and the solution left to stir at room temperature. HPLC

analysis indicated some reaction had occurred after 6 hours. Fragment 4 (**4.16**) (6.7mg, 13.7μmol, 1.eq) and DIEA (250μl, 1.45mmol) were dissolved in dry DMF (400μl) and added and the reaction left to stir at room temperature overnight. The material was desalted on a Sephadex G50 (medium grade) gel filtration column using 20% v/v AcOH/H_2O as eluent. Peptide containing fragments were pooled and lyophilised.

MS ES-MS 4865.14 (fragment 4 MW 4883), 3663.34 (fragment 5 MW 3662).

Fmoc-Glu-Val-Trp-Gln-Gly-CONHNHCONH$_2$ (4.21)

The synthesis was carried out on a 0.25mmol scale, starting from semi-carbazide linker resin (750mg, 0.36mmol/g). All amino acids were single coupled as their active HOCt ester. On completion of the synthesis the resin (850mg) was subjected to acidiolytic cleavage for 2 hours. The resin was removed by filtration and the peptide precipitated into ice cold ether, centrifuged to a pellet and washed a further 3 times with ether. The pellet was then lyophilised from MeCN:H_2O to yield the peptide as a white fluffy solid which was used without further purification.

Yield 62mg; **HPLC** (Vydac C_{18}) R_t 20.8 mins, 65% MeCN; **MS** MALDI-ToF 896.46, (MW 896.64); **Amino Acid Analysis** (24 hours hydrolysis) Glx_2 (1.96), Gly_1 (0.99), Val_1 (1.05).

H$_2$N-Leu-Ala-Leu-Leu-Ser-OH (4.22)

The synthesis was carried out on a 0.25mmol scale, starting from Fmoc-Ser(^{t}Bu) functionalised Wang resin (700mg, 0.34mmol/g). All amino acids were single coupled as their active HOCt ester. On completion of the synthesis the resin (850mg) was subjected to acidiolytic cleavage for 2 hours. The resin was removed by filtration and the peptide precipitated into ice cold ether, centrifuged to a pellet and washed a further 3 times with ether. The pellet was then lyophilised from MeCN:H_2O to yield the peptide as a white fluffy solid which was used without further purification.

Yield 140mg; **HPLC** (Vydac C_{18}) R_t 15 mins, 50% MeCN; **MS** MALDI-ToF 516.49, (MW 515.64); **Amino Acid Analysis** (24 hours hydrolysis) Ser_1 (0.87), Ala_1 (1.21), Leu_3 (2.92).

Ligated Peptide (4.23)

Semi-carbazide peptide (**4.21**) (10mg, 0.011mmol) and HOCt (235.5mg, 1.5mmol) were added to dry DMF (Aldrich) (1ml) and left to stir under an atmosphere of dry argon. ^{t}BuONO (2.6µl, 0.022mmol, 2eq) was added and the reaction left to stir at room temperature. After 2 hours an aliquot was removed and analysed by HPLC (Vydac C_{18}). After 4 hours there was no further reaction and 80% of the semi-carbazide was deemed to have converted to the active ester. Free amino peptide (**4.22**) (5.66mg, 0.011mmol, 1eq) and DIEA (384µl, 2mmol) were added and the reaction left to stir at room temperature. A sample was removed after 1 hour and analysed by HPLC, indicating all starting material had disappeared. The DMF was removed *in vacuo* and the resulting oil added to diethyl ether to precipitate the peptide. The peptide was centrifuged to a pellet and lyophilised from 50% v/v MeCN/H_2O.

Yield 2.1mg; **HPLC** (Vydac C_{18}) R_t 25.4mins, 78% MeCN; **MS** ES-MS 1337.79, (MW 1337.28); **Amino Acid Analysis** (24 hours hydrolysis) Ser_1 (0.62), Glx_2 (2.02), Gly_1 (1.01), Ala_1 (0.92), Val_1 (1.16), Leu_3 (3.23).

Synthesis of 1-Hydroxy-4-ethoxy-carbonyl-1,2,3-triazole (HOCt)

Ethyl diazoaceatate

A solution of glycine ethyl ester hydrochloride (210g, 15mol) in water (300ml) was added to DCM (500ml) and the solution cooled to –20°C (acetone/dry ice bath). A cold solution of sodium nitrite (124.5g, 18mol) in water (200ml) was added in one portion with stirring. An ice-cold solution of 5% w/w H_2SO_4/H_2O (142.5g) was added dropwise maintaining the temperature at < 15°C. The reaction was then allowed to rise to 0°C. The reaction was transferred to an ice-cold separating funnel and the yellow organic layer run in to an ice-cold solution of 5% w/v Na_2CO_3/H_2O (1l). The aqueous layer was extracted with DCM (3 x 500ml) and combined with the Na_2CO_3 layer. These were then allowed to stir until no gas evolved. The two layers were then separated and the aqueous layer extracted with DCM (500ml). The DCM extracts were combined, dried ($MgSO_4$) and the solvent removed under reduced pressure. The resulting yellow oil was stored at 5°C overnight.

The diazoacetic ester salt

Preparation of the Vilsmaier Reagent

DMF (52.8g, 0.72mol) was cooled in an ice/salt bath under an atmosphere of argon. Freshly distilled thionyl chloride (53ml, 0.72mol) was added dropwise with stirring and the solution allowed to stir for 30 mins. The resulting colourless gel was evaporated to yield the Vilsmaier reagent as a white solid.

The Vilsmaier Reagent was dissolved in chloroform (400ml) and cooled in an ice/salt bath. The ethyl diazoacetate was added dropwise keeping the temperature below 5°C. The reaction was stirred at room temp for 30 mins and the solvent removed under reduced pressure. (Due to the potential explosion hazard this was not taken to complete dryness.) Ether was added to the resulting residue to give a soft yellow solid which was filtered under nitrogen and stored in the vacuum dessicator overnight.

1-Hydroxy-4-ethoxy-carbonyl-1,2,3-triazole (HOCt)

Hydroxylamine (HCl salt, 44g, 0.63mol) was dissolved in distilled water (200ml) and cooled in an ice/salt bath. The pH of the solution was adjusted to 7-7.5 with solid Na_2CO_3. The diazonium salt (140g) was added as a solid keeping the pH at 7-7.5 by the addition of 10% Na_2CO_3. The yellow product which precipitated was stirred for 5 mins, filtered, washed with a little ice cold water, dissolved in $CHCl_3$, dried ($MgSO_4$) and filtered. Acetic acid (0.5ml) was added to this and the solution left to cyclise at 5°C. Once cyclisation occurred (confirmed by ^1H NMR) the solvent was removed under reduced pressure and the resulting oil recrystallised from EtOAc to yield *1-hydroxy-4-ethoxy-carbonyl-1,2,3-triazole* as a white solid.

Yield 21g, 21%; **Mp** 76-77°C (lit 77-79°C)[11]; **MS** FAB m/z = 158 (MH^+); **HRMS** 158.0564 (MH^+ requires 158.0566); **CHN** found C 38.27%, H 4.48%, N 26.58% ($C_5H_7N_3O_3$ requires C 38.22%, H 4.49%, N 26.74%); **^1H NMR (250MHz, $CDCl_3$ δH/ppm)** 1.33 (3H, t, J = 7.13, CH_3), 4.34 (2H, q, J = 7.13, CH_2), 8.10 (1H, s, aromatic CH), 13.96 (1H, br, OH); **^{13}C {^1H} NMR (63MHz, $CDCl_3$ δC/ppm)** (63MHz, $CDCl_3$) 13.91 (q), 61.73 (t), 121.85 (d), 136.48 (d), 159.11 (s); **FTIR ν_{max}/cm^{-1} (bromoform)** 1140 (C-O), 1536 (aromatic), 1728 (C=O), 2975 (CH), 3017 (aromatic CH), 3155 (OH); **λ_{max}/nm (MeOH, ε/dm^{3}mol^{-1}cm^{-1})** 257 (2237).

1. J Patterson, *Personal Communication*, University of Edinburgh.

2. G. L. Ellman, *Arch. Biochem. Biophys.*, 1959, **82**, 70.

3. U. K. Laemmli, *Nature*, 1970, **227**, 680.

4. F. Fülöp, J. Mattinen and K. Pihlja, *Tetrahedron*, 1990, **46**, 6545.

5. E. Brooker and U. Eisner, *J. Chem. Soc., Perkin Trans I*, 1975, 929.

6. F. Fülöp, J. Mattinen and K. Pihlja, *Tetrahedron*, 1990, **46**, 6545.

7. A. Gosden, R. Macrae and G. T. Young, *J. Chem. Res.(M).*, 1977, 0310.

8. Z. E. Zhang, Y. L. Cao, M.W. Hearn, *Anal Biochem*, 1991, **195**, 160.

9. Blicke F. F., Schaumann E. L., *J. Am. Chem. Soc.*, 1954, **76**, 3153.

10. L. C. Draffan, *Ph.D. Thesis*, University of Edinburgh, 1996.

11. L. Jiang. A. Davison, G. Tennant and R. Ramage, *Tetrahedron*, 1998, **54**, 14233.

Appendix

A.1 X-ray Crystal Data for 2-(4-pyridinyl)-1,3-thiazolidine-4-carboxylic acid (2.4)

Table 1. Crystal data and structure refinement for lb68sn.

A. CRYSTAL DATA	
Empirical formula	C9 H12 N2 O2 S
Formula weight	212.27
Wavelength	1.54184 A
Temperature	293(2) K
Crystal system	Triclinic
Space group	P1
Unit cell dimensions	a = 5.7568(10) A alpha = 75.043(10) deg. b = 8.186(2) A beta = 75.140(9) deg. c = 11.172(2) A gamma = 88.661(9) deg.
Volume	491.08(15) A^3
Number of reflections for cell	24 (20 < theta < 22 deg.)
Z	2
Density (calculated)	1.436 Mg/m^3
Absorption coefficient	2.745 mm^-1
F(000)	224
B. DATA COLLECTION	
Crystal description	Colourless block
Crystal size	0.43 x 0.16 x 0.12 mm
Theta range for data collection	4.24 to 70.17 deg.
Index ranges	-6<=h<=6, -9<=k<=9, -13<=l<=13
Reflections collected	5316
Independent reflections	3400 [R(int) = 0.0132]
Scan type	omega-theta
Absorption correction	Psi-scans (Tmin= 0.449, Tmax=0.833)
C. SOLUTION AND REFINEMENT.	
Solution	direct (SIR92)
Refinement type	Full-matrix least-squares on F^2
Program used for refinement	SHELXTL version 5
Hydrogen atom placement	Calculated and difference map (NH3+)
Hydrogen atom treatment	Riding and rotating group (NH3+)
Data / restraints / parameters	3398/3/257
Goodness-of-fit on F^2	1.006
Conventional R [F>4sigma(F)]	R1 = 0.0362 [2846 data]
R indices (all data)	R1 = 0.0455, wR2 = 0.0986
Absolute structure parameter	0.00(2)
Extinction coefficient	0.0068(11)
Final maximum delta/sigma	-0.001
Weighting scheme calc w=1/[\s^2^(Fo^2^)+(0.0655P)^2^+0.0000P] where P=(Fo^2^+2Fc^2^)/3	
Largest diff. peak and hole	0.258 and -0.258 e.A^-3

Table 2. Atomic coordinates (x 10^4) and equivalent isotropic displacement parameters (A^2 x 10^3) for 1. U(eq) is defined as one third of the trace of the orthogonalized Uij tensor.

	x	y	z	U(eq)
S(1)	6736(1)	3572(1)	-1943(1)	57(1)
C(2)	7402(4)	5580(3)	-3132(2)	35(1)
C(3)	8594(4)	5476(3)	-4494(2)	27(1)
N(3)	7141(3)	4338(3)	-4886(2)	32(1)
C(4)	11216(4)	4958(3)	-4727(2)	29(1)
O(41)	11883(3)	4059(3)	-5471(2)	44(1)
O(42)	12508(3)	5577(3)	-4175(2)	45(1)
C(13)	9268(6)	3345(5)	-1256(3)	67(1)
N(10)	8188(5)	4329(4)	2450(2)	52(1)
C(20)	10267(6)	4718(5)	1565(3)	54(1)
C(30)	10647(5)	4454(5)	358(3)	54(1)
C(40)	8838(5)	3727(4)	30(2)	42(1)
C(50)	6680(6)	3323(5)	954(3)	51(1)
C(60)	6431(6)	3642(5)	2134(3)	53(1)
S(1')	2071(1)	8498(1)	-1869(1)	50(1)
C(2')	2403(4)	10517(4)	-3052(2)	38(1)
C(3')	3543(4)	10448(3)	-4429(2)	28(1)
N(3')	2097(3)	9289(3)	-4804(2)	29(1)
C(4')	6197(4)	9977(3)	-4700(2)	28(1)
O(41')	6861(3)	9065(3)	-5432(2)	41(1)
O(42')	7493(3)	10635(3)	-4174(2)	43(1)
C(13')	4643(5)	8527(5)	-1223(3)	51(1)
N(10')	3281(5)	9206(4)	2589(2)	50(1)
C(20')	5221(6)	8471(5)	2071(3)	58(1)
C(30')	5678(5)	8239(5)	849(3)	50(1)
C(40')	4103(5)	8789(3)	108(2)	37(1)
C(50')	2090(6)	9553(5)	657(3)	49(1)
C(60')	1757(6)	9729(5)	1885(3)	51(1)

Table 3. Bond lengths [A] and angles [deg] for 1.

S(1)-C(13)	1.799(4)	C(13)-S(1)-C(2)	103.2(2)
S(1)-C(2)	1.804(3)	C(3)-C(2)-S(1)	115.1(2)
C(2)-C(3)	1.524(3)	N(3)-C(3)-C(2)	110.3(2)
C(3)-N(3)	1.482(3)	N(3)-C(3)-C(4)	110.1(2)
C(3)-C(4)	1.534(3)	C(2)-C(3)-C(4)	115.3(2)
C(4)-O(41)	1.232(3)	O(41)-C(4)-O(42)	126.7(2)
C(4)-O(42)	1.267(3)	O(41)-C(4)-C(3)	117.5(2)
C(13)-C(40)	1.506(4)	O(42)-C(4)-C(3)	115.8(2)
N(10)-C(20)	1.327(4)	C(40)-C(13)-S(1)	115.8(2)
N(10)-C(60)	1.333(4)	C(20)-N(10)-C(60)	116.9(3)
C(20)-C(30)	1.381(4)	N(10)-C(20)-C(30)	122.9(3)
C(30)-C(40)	1.382(5)	C(20)-C(30)-C(40)	120.3(3)
C(40)-C(50)	1.379(4)	C(50)-C(40)-C(30)	116.5(3)
C(50)-C(60)	1.380(4)	C(50)-C(40)-C(13)	122.5(3)
S(1')-C(2')	1.809(3)	C(30)-C(40)-C(13)	120.9(3)
S(1')-C(13')	1.809(3)	C(40)-C(50)-C(60)	119.8(3)
C(2')-C(3')	1.525(3)	N(10)-C(60)-C(50)	123.5(3)
C(3')-N(3')	1.480(3)	C(2')-S(1')-C(13')	103.08(15)
C(3')-C(4')	1.539(3)	C(3')-C(2')-S(1')	114.8(2)
C(4')-O(41')	1.230(3)	N(3')-C(3')-C(2')	110.2(2)
C(4')-O(42')	1.264(3)	N(3')-C(3')-C(4')	110.3(2)
C(13')-C(40')	1.509(3)	C(2')-C(3')-C(4')	114.8(2)
N(10')-C(60')	1.316(4)	O(41')-C(4')-O(42')	126.8(2)
N(10')-C(20')	1.327(4)	O(41')-C(4')-C(3')	117.6(2)
C(20')-C(30')	1.383(4)	O(42')-C(4')-C(3')	115.5(2)
C(30')-C(40')	1.373(4)	C(40')-C(13')-S(1')	115.6(2)
C(40')-C(50')	1.382(4)	C(60')-N(10')-C(20')	117.1(2)
C(50')-C(60')	1.381(4)	N(10')-C(20')-C(30')	123.0(3)
		C(40')-C(30')-C(20')	120.3(3)
		C(30')-C(40')-C(50')	116.2(2)
		C(30')-C(40')-C(13')	119.3(2)
		C(50')-C(40')-C(13')	124.4(2)
		C(60')-C(50')-C(40')	120.0(3)
		N(10')-C(60')-C(50')	123.4(3)

Symmetry transformations used to generate equivalent atoms:

Table 4. Anisotropic displacement parameters (A^2 x 10^3) for 1. The anisotropic displacement factor exponent takes the form: -2 pi^2 [h^2 a*^2 U11 + ... + 2 h k a* b* U12]

	U11	U22	U33	U23	U13	U12
S(1)	65(1)	73(1)	30(1)	-14(1)	-7(1)	-13(1)
C(2)	24(1)	51(2)	40(1)	-26(1)	-11(1)	12(1)
C(3)	20(1)	34(1)	32(1)	-12(1)	-9(1)	5(1)
N(3)	19(1)	49(1)	31(1)	-16(1)	-8(1)	5(1)
C(4)	18(1)	33(1)	35(1)	-7(1)	-5(1)	3(1)
O(41)	27(1)	54(1)	54(1)	-26(1)	-5(1)	11(1)
O(42)	22(1)	61(1)	58(1)	-24(1)	-14(1)	4(1)
C(13)	73(2)	99(3)	32(2)	-24(2)	-18(2)	50(2)
N(10)	44(1)	79(2)	41(1)	-28(1)	-13(1)	12(1)
C(20)	36(2)	68(2)	60(2)	-21(2)	-14(1)	2(2)
C(30)	34(2)	71(2)	44(2)	-6(2)	1(1)	7(2)
C(40)	38(2)	58(2)	28(1)	-9(1)	-8(1)	19(1)
C(50)	37(2)	80(3)	36(2)	-16(2)	-7(1)	-2(2)
C(60)	33(2)	91(3)	32(1)	-15(2)	-4(1)	1(2)
S(1')	46(1)	74(1)	29(1)	-11(1)	-9(1)	-10(1)
C(2')	23(1)	57(2)	41(1)	-29(1)	-7(1)	7(1)
C(3')	20(1)	34(1)	32(1)	-13(1)	-8(1)	6(1)
N(3')	19(1)	44(1)	29(1)	-14(1)	-9(1)	3(1)
C(4')	18(1)	33(1)	33(1)	-7(1)	-7(1)	2(1)
O(41')	25(1)	52(1)	52(1)	-25(1)	-8(1)	13(1)
O(42')	22(1)	59(1)	53(1)	-21(1)	-12(1)	2(1)
C(13')	37(2)	86(2)	34(1)	-23(2)	-8(1)	13(2)
N(10')	43(1)	76(2)	35(1)	-22(1)	-12(1)	6(1)
C(20')	42(2)	101(3)	43(2)	-30(2)	-22(1)	25(2)
C(30')	37(2)	77(2)	43(2)	-26(2)	-16(1)	22(2)
C(40')	32(1)	47(2)	30(1)	-11(1)	-7(1)	4(1)
C(50')	44(2)	70(2)	39(2)	-20(1)	-19(1)	25(2)
C(60')	41(2)	75(2)	43(2)	-27(2)	-11(1)	18(2)

Table 5. Hydrogen coordinates (x 10^4) and isotropic displacement parameters (A^2 x 10^3) for 1.

	x	y	z	U(eq)
H(2A)	5914(4)	6161(3)	-3134(2)	42
H(2B)	8448(4)	6265(3)	-2879(2)	42
H(3A)	8591(4)	6611(3)	-5059(2)	33
H(3B)	7228(29)	3274(4)	-4447(14)	48
H(3C)	7708(22)	4427(19)	-5721(4)	48
H(3D)	5615(8)	4629(16)	-4727(18)	48
H(13A)	9780(6)	2192(5)	-1168(3)	80
H(13B)	10584(6)	4088(5)	-1852(3)	80
H(20A)	11526(6)	5190(5)	1765(3)	64
H(30A)	12127(5)	4767(5)	-236(3)	64
H(50A)	5398(6)	2838(5)	782(3)	61
H(60A)	4958(6)	3361(5)	2741(3)	63
H(2'A)	827(4)	10982(4)	-3008(2)	45
H(2'B)	3378(4)	11286(4)	-2831(2)	45
H(3'A)	3482(4)	11584(3)	-4981(2)	33
H(3'1)	2160(28)	8235(4)	-4340(13)	44
H(3'2)	2685(21)	9343(18)	-5633(5)	44
H(3'3)	576(8)	9592(15)	-4665(17)	44
H(13C)	5427(5)	7463(5)	-1211(3)	62
H(13D)	5777(5)	9422(5)	-1800(3)	62
H(20B)	6324(6)	8095(5)	2550(3)	70
H(30B)	7057(5)	7709(5)	530(3)	60
H(50B)	960(6)	9948(5)	199(3)	59
H(60B)	385(6)	10245(5)	2234(3)	61

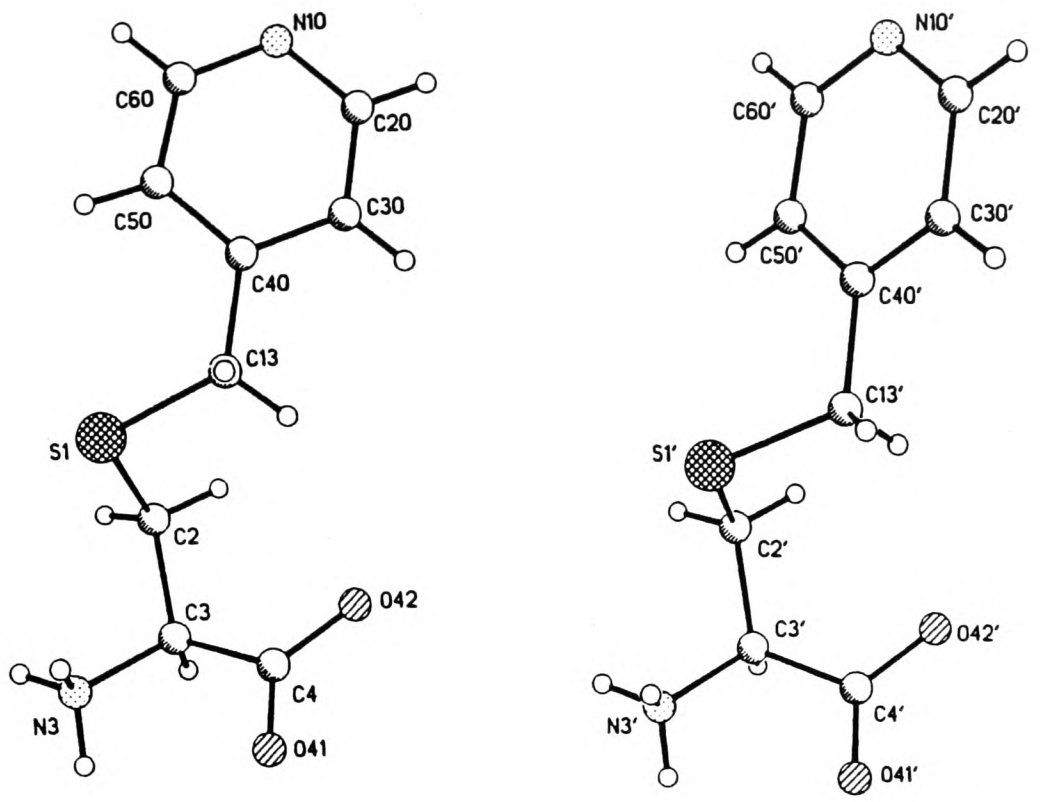
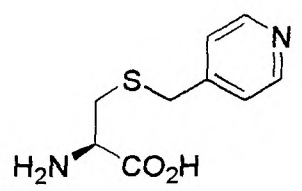


Figure A.1 X-ray Crystal Structure of 2-(4-pyridinyl)-1,3-thiazolidine-4-carboxylic acid



A.2 X-ray Crystal Data for 2-(3-pyridinyl)-1,3-thiazolidine-4-carboxylic acid (2.3)

Table 1. Crystal data and structure refinement for lb46sn at 220(2) K.

Empirical formula	C9 H12 N2 O2 S
Formula weight	212.27
Wavelength	1.54184 A
Crystal system	Monoclinic
Space group	P21
Unit cell dimensions	a = 9.934(2) A alpha = 90 deg. b = 8.777(2) A beta = 95.269(14) deg. c = 11.122(3) A gamma = 90 deg.
Volume	965.7(3) A^3
Number of reflections for cell	19 (20 < theta < 22 deg.)
Z	4
Density (calculated)	1.460 Mg/m^3
Absorption coefficient	2.792 mm^-1
F(000)	448
Crystal description	Colourless block
Crystal size	0.25 x 0.19 x 0.16 mm
Theta range for data collection	3.99 to 70.19 deg.
Index ranges	-12<=h<=12, -6<=k<=10, -13<=l<=13
Reflections collected	3843
Independent reflections	2328 [R(int) = 0.0919]
Scan type	omega-theta
Absorption correction	Psi-scans (Tmin= 0.335, Tmax=0.598)
Data / restraints / parameters	2326/1/259 (Full-matrix least-squares on F^2)
Goodness-of-fit on F^2	0.989
Conventional R [F>4sigma(F)]	R1 = 0.0587 [1606 data]
R indices (all data)	R1 = 0.1005, wR2 = 0.1359
Absolute structure parameter	0.08(6)
Extinction coefficient	0.0018(5)
Final maximum delta/sigma	0.000
Weighting scheme calc w=1/[\s^2^(Fo^2^)+(0.0597P)^2^+0.0000P] where P=(Fo^2^+2Fc^2^)/3	
Largest diff. peak and hole	0.332 and -0.374 e.A^-3

Table 2. Atomic coordinates (x 10^4) and equivalent isotropic displacement parameters (A^2 x 10^3) for 1. U(eq) is defined as one third of the trace of the orthogonalized Uij tensor.

	x	Y	z	U(eq)
S(1)	2550(2)	57(3)	2855(2)	24(1)
C(2)	3474(6)	-213(13)	4330(7)	26(2)
N(3)	5555(5)	-1212(8)	3558(5)	18(1)
C(3)	4997(6)	-64(9)	4362(6)	11(1)
C(4)	5557(7)	-277(9)	5712(7)	15(2)
O(41)	5554(5)	854(7)	6373(5)	28(1)
O(42)	5903(5)	-1592(7)	6002(5)	20(1)
C(13)	1775(7)	-1854(12)	2582(8)	24(2)
N(10)	-1521(6)	-1553(9)	4022(7)	25(2)
C(20)	-574(8)	-1487(11)	3234(9)	28(2)
C(30)	711(7)	-2101(10)	3460(7)	19(2)
C(40)	1013(8)	-2907(11)	4498(9)	25(2)
C(50)	42(8)	-2994(11)	5331(9)	28(2)
C(60)	-1176(8)	-2268(11)	5053(8)	29(2)
S(1')	2634(2)	-1895(3)	-2296(2)	24(1)
C(2')	3540(6)	-1726(12)	-810(7)	20(2)
N(3')	5821(5)	-1988(8)	-1503(5)	16(1)
C(3')	4960(7)	-1068(9)	-765(7)	16(2)
C(4')	5556(7)	-1129(9)	564(7)	15(2)
O(41')	6286(5)	-2206(7)	890(5)	30(2)
O(42')	5195(5)	-40(8)	1217(4)	28(1)
C(13')	1797(7)	-57(12)	-2414(8)	26(2)
N(10')	-1489(6)	-591(8)	-1049(7)	19(2)
C(20')	-525(7)	-584(10)	-1799(8)	20(2)
C(30')	727(6)	106(11)	-1573(7)	21(2)
C(40')	946(8)	905(11)	-489(8)	25(2)
C(50')	-33(8)	938(12)	293(9)	30(2)
C(60')	-1258(7)	175(11)	-8(8)	20(2)

Table 3. Bond lengths [A] and angles [deg] for 1.

S(1)-C(2)	1.820(8)	C(2)-S(1)-C(13)	101.3(5)
S(1)-C(13)	1.859(10)	C(3)-C(2)-S(1)	115.4(5)
C(2)-C(3)	1.516(9)	N(3)-C(3)-C(2)	110.7(6)
N(3)-C(3)	1.487(8)	N(3)-C(3)-C(4)	112.2(6)
C(3)-C(4)	1.564(10)	C(2)-C(3)-C(4)	106.2(6)
C(4)-O(41)	1.235(9)	O(41)-C(4)-O(42)	127.5(8)
C(4)-O(42)	1.239(9)	O(41)-C(4)-C(3)	117.1(7)
C(13)-C(30)	1.519(9)	O(42)-C(4)-C(3)	115.3(7)
N(10)-C(60)	1.324(12)	C(30)-C(13)-S(1)	109.0(6)
N(10)-C(20)	1.345(9)	C(60)-N(10)-C(20)	116.3(7)
C(20)-C(30)	1.387(11)	N(10)-C(20)-C(30)	123.5(9)
C(30)-C(40)	1.364(12)	C(40)-C(30)-C(20)	118.9(7)
C(40)-C(50)	1.400(11)	C(40)-C(30)-C(13)	120.4(7)
C(50)-C(60)	1.377(13)	C(20)-C(30)-C(13)	120.7(8)
S(1')-C(2')	1.813(8)	C(30)-C(40)-C(50)	118.5(8)
S(1')-C(13')	1.815(10)	C(60)-C(50)-C(40)	118.1(9)
C(2')-C(3')	1.521(9)	N(10)-C(60)-C(50)	124.5(8)
N(3')-C(3')	1.478(8)	C(2')-S(1')-C(13')	100.3(4)
C(3')-C(4')	1.542(11)	C(3')-C(2')-S(1')	116.2(5)
C(4')-O(41')	1.226(9)	N(3')-C(3')-C(2')	110.9(6)
C(4')-O(42')	1.271(9)	N(3')-C(3')-C(4')	108.8(6)
C(13')-C(30')	1.486(9)	C(2')-C(3')-C(4')	107.0(6)
N(10')-C(20')	1.327(9)	O(41')-C(4')-O(42')	126.9(8)
N(10')-C(60')	1.340(11)	O(41')-C(4')-C(3')	118.3(7)
C(20')-C(30')	1.385(10)	O(42')-C(4')-C(3')	114.7(7)
C(30')-C(40')	1.394(12)	C(30')-C(13')-S(1')	112.9(6)
C(40')-C(50')	1.364(12)	C(20')-N(10')-C(60')	117.7(7)
C(50')-C(60')	1.402(12)	N(10')-C(20')-C(30')	125.1(9)
		C(20')-C(30')-C(40')	116.3(7)
		C(20')-C(30')-C(13')	121.7(8)
		C(40')-C(30')-C(13')	121.9(7)
		C(50')-C(40')-C(30')	119.9(8)
		C(40')-C(50')-C(60')	119.4(8)
		N(10')-C(60')-C(50')	121.5(7)

Symmetry transformations used to generate equivalent atoms:

Table 4. Anisotropic displacement parameters (A^2 x 10^3) for 1.
The anisotropic displacement factor exponent takes the form:
-2 pi^2 [h^2 a*^2 U11 + ... + 2 h k a* b* U12]

	U11	U22	U33	U23	U13	U12
S(1)	16(1)	37(1)	18(1)	6(1)	-5(1)	-1(1)
C(2)	15(3)	54(7)	9(4)	3(5)	-2(3)	2(4)
N(3)	18(3)	25(4)	10(3)	-4(3)	-2(2)	5(3)
C(3)	19(3)	10(4)	3(3)	3(3)	-4(3)	2(3)
C(4)	15(3)	17(5)	12(4)	-3(4)	-4(3)	4(3)
O(41)	47(3)	18(3)	17(3)	-5(3)	-11(3)	6(3)
O(42)	32(3)	12(3)	16(3)	1(2)	-3(2)	2(2)
C(13)	28(4)	25(5)	18(4)	-9(5)	-3(4)	0(4)
N(10)	21(3)	31(5)	24(4)	3(4)	-1(3)	0(3)
C(20)	21(4)	39(7)	24(5)	0(4)	-2(4)	0(4)
C(30)	19(3)	22(5)	16(4)	0(4)	-3(3)	-5(3)
C(40)	21(4)	25(5)	28(5)	0(4)	-2(4)	2(4)
C(50)	29(4)	25(5)	31(6)	8(4)	4(4)	-6(4)
C(60)	26(4)	38(6)	22(5)	0(5)	3(4)	2(4)
S(1')	18(1)	33(1)	19(1)	-1(1)	-6(1)	2(1)
C(2')	15(3)	30(5)	13(4)	3(4)	-9(3)	0(4)
N(3')	20(3)	19(4)	9(3)	1(3)	-1(2)	-3(3)
C(3')	17(3)	11(4)	21(4)	-2(4)	-1(3)	3(3)
C(4')	22(3)	11(4)	13(4)	-4(3)	7(3)	-3(3)
O(41')	40(3)	36(4)	13(3)	-1(3)	-1(2)	21(3)
O(42')	44(3)	28(3)	12(3)	-2(3)	-1(2)	10(3)
C(13')	23(4)	31(5)	22(5)	5(5)	-5(4)	0(4)
N(10')	21(3)	25(4)	9(4)	-3(3)	-6(3)	2(3)
C(20')	21(4)	26(5)	12(5)	5(4)	-5(3)	1(3)
C(30')	15(3)	24(5)	23(5)	3(4)	-9(3)	-1(4)
C(40')	19(4)	25(5)	29(5)	2(5)	-10(4)	1(4)
C(50')	28(4)	32(6)	27(6)	-9(5)	-12(4)	-2(4)
C(60')	16(3)	25(5)	19(5)	0(4)	-4(3)	3(4)

Table 5. Hydrogen coordinates (x 10^4) and isotropic
displacement parameters (A^2 x 10^3) for 1.

	x	y	z	U(eq)
H(2A)	3147(6)	534(13)	4889(7)	31
H(2B)	3262(6)	-1228(13)	4627(7)	31
H(31)	6478(43)	-1283(39)	3737(29)	27
H(32)	5371(41)	-917(29)	2764(37)	27
H(33)	5164(38)	-2146(45)	3673(30)	27
H(3)	5227(6)	970(9)	4091(6)	13
H(13A)	1358(7)	-1919(12)	1750(8)	29
H(13B)	2472(7)	-2644(12)	2697(8)	29
H(20)	-794(8)	-998(11)	2490(9)	34
H(40)	1854(8)	-3394(11)	4650(9)	30
H(50)	216(8)	-3531(11)	6060(9)	34
H(60)	-1805(8)	-2281(11)	5635(8)	34
H(2'1)	3004(6)	-1085(12)	-312(7)	24
H(2'2)	3597(6)	-2741(12)	-441(7)	24
H(31')	5684(34)	-3037(43)	-1352(32)	24
H(32')	6741(38)	-1737(40)	-1297(31)	24
H(33')	5588(34)	-1779(40)	-2333(34)	24
H(3')	4924(7)	-1(9)	-1054(7)	20
H(13C)	2472(7)	748(12)	-2244(8)	31
H(13D)	1393(7)	84(12)	-3244(8)	31
H(20')	-706(7)	-1086(10)	-2543(8)	24
H(40')	1768(8)	1420(11)	-298(8)	30
H(50')	112(8)	1468(12)	1029(9)	36
H(60')	-1932(7)	202(11)	532(8)	24

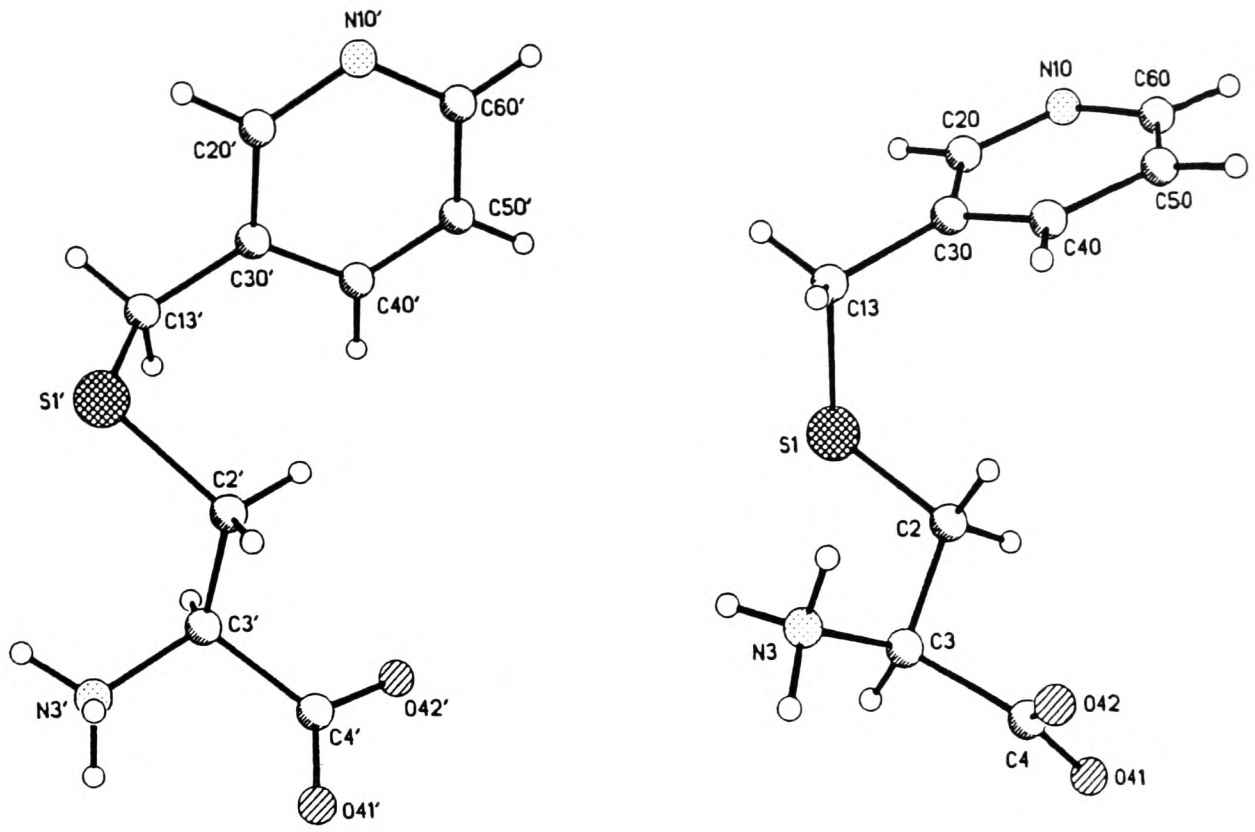
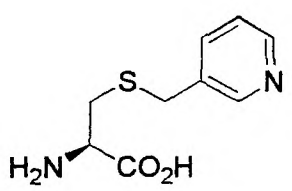


Figure A.2 X-ray Crystal Structure of 2-(3-pyridinyl)-1,3-thiazolidine-4-carboxylic acid



Courses and Conferences Attended

Departmental Colloquia, University of Edinburgh, 1996 – 1999, various speakers

Organic Research Seminars, University of Edinburgh, 1996 – 1999, various speakers

Royal Society of Chemistry, Perkin Division, Scottish Meeting, University of Edinburgh 1996, Strathclyde University 1997, St. Andrews University 1998, various speakers

SCI Graduate Symposium, Novel Organic Chemistry, University of Edinburgh 1998, University of Glasgow 1999, various speakers

5th Solid Phase Synthesis and Combinatorial Libraries Symposium, London, September 1997, various speaker

6th Solid Phase Synthesis and Combinatorial Libraries Symposium, York, September 1999, various speakers

SCI One Day Syposium, Understanding Protein Protein Interactions, London, June 1998, various speakers

RSC Conference, Biological Challenges in Synthetic Organic Chemistry, St. Andrews University, July 1999, various speakers

Current Awareness in Organic Chemistry, (sponsored by Zeneca Grangemouth), University of Edinburgh, May 1997, 1998, 1999, various speakers

Edinburgh Centre for Protein Technology lectures, University of Edinburgh, 1996 – 1999, various speakers

Walker Memorial Lectures, University of Edinburgh, 1996 – 1999, various speakers

RSC Conference, Medicinal Chemistry, Edinburgh, September 1998, various speakers

Centre for Teaching and Learning Assessment, Demonstrating Course, University of Edinburgh, October 1996